

The Unintended Consequences of School Curricula Reforms: Examining Spillover Effects of US Education Standards *

Benjamin W. Arold & M. Danish Shakeel

April 5, 2023

Abstract

School curricula have lasting impacts on students. In theory, reforming school curricula can be a powerful tool for policy-makers, but may also be harmful. To empirically prove the latter, we analyse the adoption of the Common Core State Standards (CCSS) for math and English in the US. We estimate spillovers of the CCSS on student achievement in non-targeted subjects such as science and social studies in models with state and year fixed effects. Using individual-level student achievement data from NAEP, we show that the CCSS had a negative effect on student achievement in non-targeted subjects. This effect is largest for young underprivileged students. To uncover mechanisms, we use teacher surveys to show that the CCSS reduced the instructional focus on non-targeted subjects. Our results call for a careful evaluation of potential spillovers on non-targeted outcomes when designing targeted education reforms.

JEL-Codes: H75; I21; I24; I28; J24

Keywords: School Curricula, Common Core, Education Standards, Student Achievement, Education Policy

* **Arold** (corresponding author): ETH Zurich (aroldb@ethz.ch); **Shakeel**: University of Buckingham (danish.shakeel@buckingham.ac.uk). This paper was previously circulated under the title "The Unintended Effects of the Common Core State Standards on Non-Targeted Subjects". This paper has greatly benefited from discussions with Joshua Bleiberg, Davide Cantoni, David Figlio, Leslie Finger, Eric Hanushek, David Houston, Melissa Lyon, Paul Peterson, Martin West, Ludger Woessmann, and Larissa Zierow. We are also grateful to seminar participants at Harvard, LMU Munich, and the ifo Institute as well as the Society of Labor Economists (SOLE) 2021 Conference, the Royal Economic Society (RES) Annual Conference 2021, the Annual Meeting of the American Educational Research Association (AERA) 2021, the Association for Education Finance and Policy (AEFP) Annual Conferences 2020 and 2021, as well as the American Political Science Association (APSA) Meeting 2020. Antonio Wendland provided outstanding administrative support. Laure Fleury provided excellent research assistance. Arold thanks Paul Peterson and the Program on Education Policy and Governance at Harvard University for their hospitality while writing parts of this paper, and gratefully acknowledges financial support by the DAAD through a one-year scholarship for doctoral students funding his stay at Harvard University. Financial support by the Leibniz Competition (SAW 2019) is also gratefully acknowledged. All errors are our own.

1 INTRODUCTION

Policies can have unintended effects. In the economics of crime, it has been shown that criminalizing sex work can increase sexually transmitted infections among sex workers (Cameron et al., 2020). Research in development economics documents that gold-mining activities can increase malaria incidence (Rozo, 2020). In the economics of education, it has been demonstrated that affirmative action policies can lower educational outcomes for students from low-SES neighborhoods (Barrow et al., 2020). This recent wave of empirical papers documents far-reaching unintended consequences of policies in various areas and calls for a comprehensive assessment of potential policy impacts. We bring this research agenda to the area of school curricula and ask whether targeted reforms of school curricula can have unintended consequences too.

School curricula have lasting impacts on students. They affect students' skills and productivity (Cortes and Goodman, 2014; Fuchs-Schündeln and Masella, 2016; Goodman, 2019; Conger et al., 2021), as well as their identities, preferences, and beliefs (Bernheim et al., 2001; Clots-Figueras and Masella, 2013; Cantoni et al., 2017; Bazzi et al., 2020). Recent evidence has shown that the topical content of education standards has lasting effects on students' related knowledge, beliefs, and choices (Arold, 2022). Hence, school curricula in general, and the content of education standards in particular, can be powerful tools for policy-makers to achieve their goals (Lott, 1999; Cantoni and Yuchtman, 2013). However, reforms of the "content of education" can also backfire. In theory, such reforms can (i) be counterproductive for achieving the policy goal in question, or (ii) cause negative spillovers on other policy goals. Fouka (2019) demonstrates the former by showing that language restrictions for migrants in elementary school aimed at integration negatively affect identification with the host country later in life. Our paper demonstrates the latter by showing that reforms of the content of education can have negative spillovers on non-targeted outcomes.

Specifically, we study the centralization and modernization of US education standards in math and English and test how it affected student achievement in other school

subjects. We evaluate the adoption of the US "Common Core State Standards" (CCSS), an internationally benchmarked education standard (Jerald, 2008) in math and English Language Arts (ELA), which does not address other subjects such as science and social studies. Our focus on the CCSS to study spillovers of targeted school curricula is inspired by anecdotal evidence of teachers suggesting the CCSS reduced the instructional focus on non-targeted subjects. The empirical literature on the effects of the CCSS on student achievement has so far documented zero to modest positive effects on student achievement in the targeted subjects: math and ELA.¹ In our paper, we also estimate the effect of the CCSS on student achievement in targeted subjects and, again, find zero to modest positive effects. This finding largely confirms the conclusions of the prior literature, although our robustness checks indicate that prior analyses rather overestimate any positive effects, see Appendix B for more details. However, the main focus of our paper is on spillovers of the CCSS on student achievement in non-targeted subjects such as science and social studies. We argue that studying such spillovers in a causal framework is essential for evaluating the overall success of the CCSS in terms of student achievement, and for guiding future reforms of education standards and school curricula in general.

In theory, it is unclear whether any spillovers of the CCSS on student achievement in non-targeted subjects are positive or negative (or zero). On the one hand, the CCSS could be beneficial beyond its target subjects if, for example, skills acquired in a targeted subject such as math help students to perform well in a non-targeted subject such as science. On the other hand, the CCSS could have caused a reduction of instructional focus on the non-targeted subjects, possibly leading to a decline in student achievement in those subjects.

¹Most of the early studies are correlational and find zero to modest positive associations between CCSS and student achievement in targeted subjects (Schmidt and Houang, 2012; Loveless, 2014, 2015, 2016). Exploiting quasi-experimental settings, more recent papers confirm the zero to modest positive effects (depending on study and subgroups therein) for Kentucky (Xu and Cepa, 2018), California (Gao and Lafortune, 2019), Chicago (Allensworth et al., 2021) and a subset of US states (Bleiberg, 2021). An exception is Song et al. (2019) who find negative effects on student achievement in targeted subjects of the adoption of general College and Career Readiness Content Standards (CCRCS). This study cannot be directly compared to the other studies, as the CCRCS include the CCSS but also other education standards.

Simple correlations between CCSS adoption and student achievement in non-targeted subjects likely do not yield a causal answer to our research question. States that adopted the CCSS plausibly differ from states that did not adopt the CCSS in ways that affect student achievement through many channels other than education standards, for example through differences in political preferences or human capital. To overcome this identification problem, we estimate the effect of the CCSS on student achievement in non-targeted subjects in a two-way fixed effects difference-in-differences (DD) framework. This approach builds on the idea that states without CCSS adoption in a certain year act as counterfactuals for states with CCSS adoption in that year, after accounting for time-invariant differences between states and national differences between years.

To run these DD models, we combine state-level data on the adoption of the CCSS with individual-level student achievement data from the National Assessment of Educational Progress (NAEP). The NAEP is known as The Nation's Report Card and provides unique student-level test score data for a large number of years, grades and subjects. In particular, it covers a range of subjects not targeted by the CCSS, namely science and different social studies (civics, economics, geography, and history). It is comparable across states and over time and covers the relevant years before and after the adoption of the CCSS. The NAEP student and teacher surveys complement the test score data by providing information on student characteristics, teacher characteristics and classroom instruction.

We find a significant negative effect of the CCSS on student achievement in non-targeted subjects. More specifically, being exposed to the CCSS for the entire school career (at the time of testing) as opposed to not being exposed to the CCSS at all decreases student achievement in non-targeted subjects on average by 0.08 units of a standard deviation. The effect size can be interpreted as a loss of learning worth approximately 25 percent to 30 percent of a school year. We regard this finding as reduced-form evidence that the CCSS induced a reduction of instructional focus on non-targeted subjects.

Next, we hypothesize that the negative effect is over-proportionally large for underprivileged students as these students (or their parents and environments in general)

might be less able to compensate for the reduction of instructional focus in non-targeted subjects. To test this hypothesis, we conduct subgroup analysis by student characteristics and find that the negative effect is mostly driven by Black and Hispanic students, and by students with free or reduced price lunch status, English language learner status, and disability status. We also conduct subgroup analyses by subject and grades and find that the achievement losses are most pronounced for science and for students in grade 4.² In sum, we conclude that the CCSS mainly reduced student achievement in non-targeted subjects among young underprivileged students, exacerbating racial/ethnic and socioeconomic student achievement gaps.

A series of robustness checks supports our main results. We show that the results are robust to event-study specifications and specifications with state-specific time trends. To account for state-specific shocks simultaneous to the adoption of the CCSS, we run a triple-difference model with students from private schools as an additional control group for which the CCSS was never mandatory. We also control explicitly for policy variables such as the party of the ruling governor, and a wide range of educational reforms. Furthermore, we account for recent developments in the econometric literature on two-way-fixed effects models and time-varying treatment effects and show that our results are not driven by negative weights (Chaisemartin and D’Haultfoeuille, 2020; Borusyak et al., 2021; Callaway and Sant’Anna, 2021; Goodman-Bacon, 2021; Sun and Abraham, 2021). Another set of robustness checks defines treatment based on information about actual CCSS implementation in the different states to account for the fact that CCSS adoption and CCSS implementation could diverge.

To uncover the mechanisms behind these results, we aim to understand what has changed in the classrooms of the students due to the CCSS. To this end, we draw on the NAEP teacher survey data on instructional focus. We find that the CCSS reduced teacher-reported instruction time, instructional resources, and some dimensions of the quality of

²In addition to the CCSS, some states adopted the Next Generation Science Standards (NGSS). The NGSS were released in 2013 with slow take-up rates by states relative to the CCSS. In appendix A, we demonstrate that our results are robust to controlling for NGSS adoption.

teacher-student interactions for the non-targeted subjects. This finding suggests that the exclusion of science and social studies from the CCSS has signaled a lower relative importance of these subjects, resulting in a reduction of instructional focus.

The results presented in this paper extend the literature on strategic reactions of educators to reforms to the topic of education standards. While there is evidence that accountability reforms such as the No Child Left Behind Act (NCLB) can shift the instructional focus away from subjects, topics, and students not targeted by the accountability rating (Jacob, 2005; Reback, 2008; Neal and Schanzenbach, 2010), little is known about analogous effects of reforms of education standards. The CCSS is conceptually linked to NCLB, but focuses on setting standards rather than incentivizing accountability.³ Analogous to our findings on the CCSS, NCLB reduced instruction time in non-targeted science and social studies (Reback et al., 2014).⁴ However, while we find for the CCSS that the reduction in instructional focus on the non-targeted subjects coincides with student achievement declines in these subjects, NCLB and other accountability reforms did not (necessarily) cause equivalent achievement declines in the non-targeted subjects (Winters et al., 2010; Dee and Jacob, 2011; Reback et al., 2014). One possible explanation for this difference is that the CCSS has led to much less pronounced student achievement gains in the targeted subjects relative to NCLB. This lack of substantial student achievement gains in the targeted subjects under CCSS may have limited positive spillovers on student achievement in non-targeted subjects, and may have made shifts in instructional focus matter more.

³NCLB was an accountability reform from 2002 that compelled states to design school accountability systems based on annual student assessments in math and ELA that were linked to state standards. To help states improve and meet the NCLB standards, Common Core provides education standards in math and ELA defining the knowledge and skills expected of students at each grade level. Furthermore, states that adopted the CCSS could get a waiver from some of the NCLB regulations. In this sense, our estimated effect of Common Core adoption combines both the impact of introducing Common Core standards and removing part of NCLB. Additional analyses presented in Section 6 suggest that the overall effect is not primarily coming from the NCLB waivers. In 2015, the Every Student Succeeds Act (ESSA) replaced NCLB and shifted its federal accountability aspect to the states.

⁴Anecdotal evidence attributes this to NCLB's measure of adequate yearly progress (AYP) for schools which has been determined through annual growth in reading and mathematics. As a result, instruction time and other dimensions of instructional focus such as teacher practice support have been reduced for science and social studies (see Settlage and Meadows (2002) and Johnson (2007) for science, and Misco (2005) and National Council for Social Studies (2008) for social studies).

Our main outcome variable student achievement is not only interesting in its own right, but also an important predictor of economic outcomes at the individual and societal level. Student achievement has been found to affect earnings, income distribution, and economic growth (Hanushek and Woessmann, 2008, 2012). Notably, the predictive power of student achievement is much stronger than that of traditional measures of human capital used in the literature such as literacy rates (Romer, 1990), school enrollment (Barro, 1991), or years of education (Barro and Lee, 2013). And although student achievement does not adequately capture non-cognitive skills (Heckman and Kautz, 2012; Jackson, 2018), student achievement has also been shown to be a strong predictor of cognitive skills, physical and mental health, and voting behavior (Borghans et al., 2016).

The paper proceeds as follows. Section 2 provides institutional background about the adoption and implementation of the CCSS. Section 3 outlines the empirical approach, while Section 4 describes the data. Section 5 reports the main results and heterogeneities of the effect of CCSS exposure on student achievement in non-targeted subjects. A series of robustness checks is presented in Section 6. Section 7 shows additional analyses of mechanisms, and Section 8 concludes.

2 INSTITUTIONAL BACKGROUND

2.1 Background and Data on the Adoption of the CCSS

Although US student achievement has risen substantially over a half century, it has been lagging behind the student achievement of many other industrialized countries for a number of decades (Hanushek et al., 2012; Shakeel and Peterson, 2022). The adoption of rigorous centralized education standards as a means of aligning basic elements of local school curricula has long been proposed to raise US student achievement (Costrell, 1994; Bishop, 1997; Costrell, 1997; Betts, 1998). In 2008, the National Governors Association (NGA), the Council of Chief State School Officers (CCSSO), and Achieve Inc. jointly published a report titled “Benchmarking for Success: Ensuring U.S. Students Receive

a World-Class Education.” (Jerald, 2008). The report prescribed that US states adopt a common core of internationally benchmarked standards in math and ELA to raise US achievement levels on international assessments. A number of philanthropic organizations, including the Bill & Melinda Gates Foundation, provided resources to enable the states to establish a common core of standards. Subsequently, numerous state governments, teachers’ unions and other interest groups advocated for a systemic change in education standards across the nation. In 2009, the U.S. Department of Education set incentives for states to adopt the CCSS. If a state adopted the CCSS, it could get a waiver from some of the No Child Left Behind (NCLB) regulations.

Our primary source for state-level data on the adoption of the CCSS is Achieve Inc. (2013), with an updated version provided by CCSSI (2023). CCSSI (2023) is the website of the Common Core State Standards Initiative provided by the National Governors Association Center for Best Practices and the Council of Chief State School Officers. They report if and when a state has adopted the CCSS.⁵ Based on this data source, 42 states have adopted the CCSS permanently. Of the states that have adopted the CCSS permanently, most states adopted it in 2010, while a number of states adopted it in 2011 and one state, Wyoming, adopted it in 2012. In contrast, Alaska, Florida, Indiana, Nebraska, Oklahoma, South Carolina, Virginia and Texas have not adopted the CCSS permanently. In our baseline coding, we code students in these 8 states as never having been treated by the CCSS.⁶ To account for the fact that some of those 8 states had adopted the CCSS temporarily, we present robustness checks in which we treat temporary adopters as treated from the year of the temporary adoption, even if they repealed/revised the CCSS later on. In the latter coding, only the states that never adopted the CCSS, Alaska, Nebraska, Virginia and Texas, remain in the control group, based on data from Bleiberg (2021) and CCSSI (2023). The map presented in Figure 1 illustrates the CCSS adoption graphically.

⁵Throughout the paper, we treat Washington D.C. as a state.

⁶Minnesota is a special case as it did not adopt the CCSS for math, but for ELA, which we code accordingly.

2.2 Background and Data on the Implementation of the CCSS

The implementation of the CCSS was not straightforward. There is anecdotal evidence that the CCSS presented challenges in teaching and testing to schools. Some teachers had difficulty adjusting to the new curriculum, and CCSS-based standardized tests were not always appropriate for students' ability levels. A case study in New York found that the CCSS led to exceedingly long and difficult exams. The rigor of the standardized tests exceeded the level of college readiness and represented more of an early college level (Polleck and Jeffery, 2017). In addition, some states did not have assessments and textbooks aligned with the CCSS until 2013 or later (Polikoff, 2017). Although not all challenges of the implementation of the CCSS have been overcome everywhere (Polikoff, 2015; Bay-Williams, 2016) and in some contexts teachers do not have standards-aligned materials (Polikoff et al., 2020), surveys show that many teachers feel prepared to teach the CCSS (Scholastic, 2014), have acquired good or excellent knowledge of the CCSS (Kane et al., 2016), base their curricula on the CCSS (Opfer et al., 2016), and use textbooks based on the CCSS (Blazar et al., 2019). Still, the challenges surrounding the implementation of the CCSS warrant robustness checks of the treatment coding in which we base the definition of students' CCSS exposure on the implementation of the CCSS, not just its adoption.

To account for this issue, we draw on a variety of data sources to create alternative treatment indicators based on CCSS implementation. They incorporate information on states' legal CCSS implementation requirements, actual CCSS implementation strategies, effectiveness of CCSS implementation, temporal CCSS implementation, and CCSS-aligned standardized testing. Detailed explanations of the different treatment indicators and their data sources are provided when reporting the corresponding results in Section 6 on robustness.

3 IDENTIFICATION STRATEGY

To estimate the effect of the CCSS on student achievement in non-targeted subjects, we run a two-way fixed effects difference-in-differences model (DD), and several extensions including models with state-specific time trends and a triple-difference model (DDD). The DD model takes advantage of the fact that some states did not adopt the CCSS. This approach builds on the idea that states without reforms in a given year act as counterfactuals for states with reforms in that year, after accounting for time-invariant differences between states and national differences between years. To capture this idea econometrically, we estimate a DD model as follows:

$$T_{istuv} = \beta * CCSS_Exposure_{istuv} + \gamma * X_{istuv} + \mu_s + \lambda_t + \theta_u + \kappa_v + \epsilon_{istuv} \quad (1)$$

where T_{istuv} captures standardized student achievement of student i who goes to public school in state s , and takes the test in year t , grade u and subject v . Our main estimates pool all subjects that are not targeted by the CCSS, across all available grade levels. The treatment parameter $CCSS_Exposure_{istuv}$ captures the CCSS exposure of student i attending public school in state s , and taking the test in year t , grade u and subject v . It is a dosage parameter, defined as the share of schooling years in which a student was exposed to the CCSS. β is the parameter of interest capturing the effect on student achievement of being exposed to the CCSS for the entire school career until the test date (exposure=1) relative to never being exposed to the CCSS until the test date (exposure=0).⁷ In our preferred treatment coding, we define a year in a given state as exposing a student to the CCSS if the state had adopted the CCSS permanently before that year or in the same year. In robustness checks, we employ other treatment definitions.

⁷The cross-sectional data structure of NAEP only allows us to observe students at the test date, but not to follow them through their school career. Hence, we have to assume that students did not move between states during school age before the test. To the extent that pre-test moving patterns between states are uncorrelated with CCSS adoption, this merely causes attenuation bias. In that sense, our effects should be interpreted as lower-bound estimates.

A vector of student-level control variables X_{istuv} includes indicator variables for gender, race/ethnicity, subsidized lunch status (indicator variable equals one if student receives free or reduced price lunch), English language learner status, disability status, parental education, and home possessions (separate indicator variables for computer and books). State fixed effects μ_s , test year fixed effects λ_t , grade fixed effects θ_u , subject fixed effects κ_v , and an error term ϵ_{istuv} complete the model. Note that test year and grade jointly define each cohort. Throughout the paper, all standard errors are clustered at the state level to account for potential correlation of error terms across years within states. Regressions are weighted to be population representative. We run the main DD estimations on a sample of students attending public schools (district and charter schools) only, as the implementation of the CCSS was never mandatory for private schools.

This baseline model addresses a variety of concerns about our ability to estimate the causal effect of CCSS exposure on student achievement in non-targeted subjects. First, one might be worried that state-level differences in domains such as returns to education, cultural characteristics that promote educational success, genetic endowments, or preferences for centralizing policies are correlated with CCSS exposure and affect student achievement. The state fixed effects eliminate all constant differences between states. Hence, we exploit cross-cohort variation within states. Second, one might be concerned that national trends in student achievement, for example fueled by overall economic development or national education policies, appear as effects of CCSS exposure. However, our year fixed effects capture all variation in student achievement that occurs nationwide between years. In addition, our individual-level control variables ensure that the students we compare are similar with regards to demographic and socioeconomic characteristics. For these reasons, our DD model yields a causal effect of CCSS exposure on student achievement in non-targeted subjects if the main identifying assumption about parallel trends holds. It assumes that in the absence of states adopting the CCSS, the change in student achievement in treated states would have been the same as that in non-treated states.

Although this assumption cannot be directly tested, we perform a series of robustness checks to assess its plausibility. We begin with running non-parametric event-study specifications, in which the adoption of the CCSS in a given state and year is defined as the event. In contrast to the DD model, the event-study model can assess non-linear pre-reform trends in student achievement. If student achievement prior to the adoption of the CCSS was trending in the direction of the estimated CCSS effects, this could indicate a bias from underlying trends in the data. Another advantage of the event-study model is that the time course of effects of the adoption of the CCSS can be assessed by disentangling effects which occur directly at the time of the CCSS adoption from those which occur gradually after the CCSS adoption. Specifically, we estimate the effect of the CCSS adoption in year t_s on student achievement k years before and after CCSS adoption, as captured by the parameter vector β_k , see equation (2). These effects are estimated relative to the year of reform $k=0$.⁸

$$T_{istuv} = \sum_{k=-6}^6 1(t_{is} = t_s + k)\beta_k + \gamma * X_{istuv} + \mu_s + \lambda_t + \theta_u + \kappa_v + \epsilon_{istuv} \quad (2)$$

Although the non-parametric specification captures an overall pre-trend, it does not account for state-specific trends. To address this issue, we perform analyses in which we add state-specific linear time trends, linear and quadratic time trends, as well as linear, quadratic and cubic time trends to equation (1). The state-specific linear time trend variable interacts each state fixed effect with a re-scaled year variable that equals one in the first year of observation, two in the second year of observation, and so forth. The corresponding reform effect is identified from within-state deviations in student achievement from smooth linear trends that coincide with the different timing of CCSS adoption across states. State-specific quadratic and cubic time trends are defined analogously. Taken together, the event-study model and the models with state-specific trends address concerns about underlying trends in the data.

⁸To smooth the numbers of observations across years, the observations are grouped together to bins of 2 years for all pre- and post reform years except for the bins at the beginning (end) of the domain which additionally include the years prior to (following) the domain's starting (ending) year.

Even if there are no underlying trends in the data, shocks or events that occur simultaneously to the adoption of the CCSS remain a threat to the parallel trends assumption. To address this issue, we first run a triple difference model (DDD). For this analysis, we add an additional control group of private school students to our DD sample of public school students. The CCSS has never been mandatory for private schools. Correspondingly, we code private school students as not being exposed to the CCSS, even if their school is located in a state that has adopted the CCSS for public schools in some years of the school career of the students. Given the possibility that some private schools might have voluntarily implemented some elements of the CCSS, we should interpret DDD effects as lower-bound estimates,⁹ (at least under the assumption of no endogenous selection between public and private school students as we discuss below).

Econometrically, we capture our third difference using a school type indicator variable (public school vs. private school). The full DDD follows equation (1) but adds a baseline indicator for school type as well as a full set of fixed effects interactions. This set includes school-type-by-state fixed effects (for example to control for state-specific time-constant regulation differences between private and public schools), school-type-by-year fixed effects (for example to control for changes in the national funding of public schools), and state-by-year fixed effects (for example to control for state-specific policies and programs directed at students or their families regardless of school type). The DDD uses variation at the school-type-by-state-by-year level to identify the effect of the CCSS on student achievement from differences in student achievement of students who attend public school compared to student achievement of students who attend private school coincident with the timing of the CCSS adoption in each state. The identifying assumption of the DDD requires that there is no other school type specific variable correlated with the CCSS adoption that affects student achievement. This identifying assumption is substantially weaker than that of the DD model, as it cancels out all confounding variables

⁹If some private schools implemented elements of the CCSS, we would code some students as untreated although they have in fact received some treatment. Hence, we would erroneously difference out some part of the real effect of the CCSS.

that affect public and private school students equally.

Still, policies that occurred simultaneously to the adoption of the CCSS which affect public and private school students differently and which have an effect on student achievement could bias the models presented so far. To address this concern, we collect data on reforms of public schooling policies and private schooling policies. Examples of public schooling policy controls include waivers from NCLB/ESSA accountability requirements, the adoption of the Next Generation Science Standards, and public education expenditures (as measured by the district-by-year-level per-pupil education expenditures in logarithmized dollars). Examples of private schooling policy controls include states' control of private school licensure, of private school curricula, or publicly funded voucher laws. Similarly, we also add controls for policies on homeschooling and compulsory schooling as a robustness check, as they might indirectly and differently affect public and private schools. The extent of recordkeeping requirements for homeschooling is an example of a homeschooling policy, and the number of compulsory schooling years is an example of a compulsory schooling policy. Adding these policy variables as controls can alleviate many concerns about simultaneous policies biasing the results. In a similar vein, we also explicitly control for the political party of both the state governor in power and the state legislative chamber.

Another threat to validity is that the adoption of the CCSS might have caused heterogeneous selection of specific groups of students into school types. For example, estimates could be biased if students with politically conservative parents left the public school system at the same time that the CCSS was adopted, and if political conservatism of parents affects student achievement. Neither the DDD model, nor the DD models based on a sample of public school students, are immune to this selection issue. We address this concern by running a DD model with a joint sample of public and private school students, whose reform effects are net of any heterogeneous selection between school types.

Even if the parallel trends assumption holds, the previously presented two-way fixed effects models could yield biased results due to time-varying treatment effects. If already-

treated students act as controls for later-treated students in settings with staggered treatment timing, time-varying treatment effects can bias results away from the true effect. This issue, also referred to as "negative weighting", has received much attention in the recent econometric literature (Chaisemartin and D'Haultfœuille, 2020; Borusyak et al., 2021; Callaway and Sant'Anna, 2021; Goodman-Bacon, 2021; Sun and Abraham, 2021). We address this issue by conducting a robustness check in which we exclude all 2x2 DD comparisons from the sample in which already-treated students act as controls. In sum, each of the presented approaches in this section has different identifying assumptions and addresses concerns about underlying trends in the data, simultaneous shocks, selection, or negative weights, in a different way. In our view, robust insights about the effect of the CCSS on student achievement in non-targeted subjects can be obtained if the different approaches yield similar results.

4 DATA

4.1 *Student Achievement Data*

We merge data on the adoption and implementation of the CCSS with standardized student achievement data. The data sources on the adoption and implementation of the CCSS are described in Section 2 on the institutional background and in Appendix C. For student achievement, we use the restricted-use individual-level dataset of the NAEP. The NAEP is a congressionally mandated project which is nationally representative of the US student body. It is also known as The Nation's Report Card. NAEP has measured the knowledge of US students in various subjects since its first assessment was conducted in 1969.¹⁰ The assessments are administered by the National Center for Education Statistics (NCES), an institution within the Institute of Education Sciences (IES) and the US Department of Education. Notably, there is a significant but not complete overlap between CCSS and NAEP items (Daro et al., 2015). For example, 79 percent of items on grade 4

¹⁰Throughout the paper we use the Main-NAEP and not the Long Term Trend NAEP, as the Main-NAEP has much larger sample sizes and is state-representative.

and 87 percent of items on grade 8 of the 2015 NAEP math assessment match the content that is present in CCSS at or below those grade levels.

The NAEP individual-level dataset has several advantages for our analysis. First, it provides information on student achievement at the individual level for the relevant years, grades, and subjects for all US states. Its comparability across states and over time allows for consistent standardization and two-way fixed effects difference-in-differences estimations. Specifically, we use the NAEP data from 2005 to 2015 for all available grade levels, namely grades 4, 8 and 12.¹¹ We exclude data from 2004 and before, as NAEP sampling increased tremendously after 2001 (and no testing was done in non-targeted subjects between 2002 and 2004 and after 2015). Besides, this sample cut provides pre- and post reform periods of roughly equal duration. We use the NAEP data from science, civics, economics, geography, and history to capture student achievement in non-targeted subjects.¹² Table A.1 lists the grades in which the NAEP tests were administered for each of the five subjects for each year between 2005 and 2015 (and for which state identifiers of students are available). We use the student-level data from all these subject-year-grade combinations in our main analysis. The resulting sample consists of more than one million students.^{13 14}

Second, the NAEP data include a rich set of individual-level control variables such as student gender, race/ethnicity, and various socio-economic background variables, among others. Hence, we can control flexibly for students' pre-reform demographics and perform subgroup analyses. We set missing values of controls to zero and add separate explanatory binary variables to all regressions to account for these missing values, unless noted

¹¹For the targeted subjects evaluated in Appendix B, we have data until 2017.

¹²Other tests such as theater, visual arts, and music were not tested in the relevant years.

¹³Throughout the paper, we report the number of observations rounded to the nearest ten digit to comply with data protection regulations of the NCES.

¹⁴In the NAEP, no student takes the entire student achievement test. Instead, the NAEP reports plausible values for overall student achievement on a test estimated from the sample of questions that were administered to a student. We make the arbitrary choice of selecting the second plausible value the NAEP provides. The raw correlation between the second plausible value and the average of the first five plausible for student achievement in, for example, science equals 0.95 in our sample. Our results are robust to estimating effects using any other plausible value.

otherwise.¹⁵ Third, the NAEP also administers student achievement tests and surveys in private schools, which we can exploit for identification given that the CCSS was never mandatory for private schools. Fourth, the NAEP has also surveyed the teachers of a subset of the students in the sample. We can make use of the teacher data to investigate changes in classroom instruction resulting from the CCSS, as presented in Section 7 on mechanisms.

4.2 *Descriptive Statistics*

Table A.2 presents the mean, standard deviation, minimum and maximum of the main variables. The main outcome variable is student achievement, which we standardize to have a mean of zero and a standard deviation of one in the first year of each available grade-subject combination, following Lafortune et al. (2018) and Jackson et al. (2021). This standardization allows the mean and variance to flexibly evolve in the following years, which explains why their overall mean and standard deviation reported in Table A.2 are close but not equal to zero and one, respectively. Regarding student characteristics, we note that about half of the sample is female and almost 60 percent is White. The shares of Black and Hispanic students are 15 percent and 20 percent, respectively. 5 percent of the students in the sample are Asian students (including Pacific Islanders). About 6 percent of the sample have English language learner status, and 11 percent disability status. To assess what share of students come from a low socioeconomic background, we can look at the shares of students receiving subsidized lunch (43 percent), having parents who did not finish high school (8 percent), having no computer at home (10 percent), or having less than 10 books at home (13 percent). We also show descriptive statistics for variables measuring the instructional focus on non-targeted subjects and teacher characteristics (both based on the NAEP teacher surveys), which we analyze in more detail in Section 7.

¹⁵Our results are robust to not imputing the missings, see Table A.10.

5 RESULTS

5.1 *Main Results*

Table 1 presents estimates of the statistical relation between CCSS exposure and student achievement in non-targeted subjects for different sets of control variables. In column 1, there are no control variables. The positive and statistically significant correlation between CCSS exposure and student achievement in non-targeted subjects implies that students exposed to the CCSS perform better in non-targeted subjects than those not exposed to the CCSS. This correlation could be caused by CCSS exposure improving student achievement in non-targeted subjects, for example through positive spillovers. It could also be caused by above-average student achievement in non-targeted subjects leading to CCSS exposure. This reverse causality could occur, for example, if above-average student achievement in non-targeted subjects before the reform promotes confidence in national education policies, thereby encouraging states to adopt the CCSS. Moreover, the observed positive correlation between CCSS exposure and student achievement in non-targeted subjects could also be driven by third variables such as parental education. This would be the case, for example, if states with a high proportion of students with highly educated parents are more likely to adopt the CCSS, for instance if these parents vote disproportionately for parties that endorse the CCSS, and if parental education itself increases student achievement in non-targeted subjects.

To isolate the effect of CCSS exposure on student achievement, we add control variables in columns 2-3. The positive correlation of CCSS exposure and student achievement in CCSS subjects remains almost unchanged conditional on student-level control variables (column 2). In the full model with both student-level controls as well as state and year fixed effects, the positive correlation becomes negative (column 3). The full model is our preferred model, as it flexibly accounts for demographic and socioeconomic differences between students as well as time-invariant differences between states and national differences between years. Unless noted otherwise, all further models

presented in this paper are full models.

More specifically, exposure to the CCSS during the entire school career, as opposed to no CCSS exposure at all, decreases student achievement in non-targeted subjects on average by 0.08 units of a standard deviation. This effect is statistically significant at the 5 percent level. To illustrate the effect size, we draw on the literature on education production functions, which suggests that the gain in learning from one year of schooling is equivalent to about one-quarter to one-third of a standard deviation increase in student performance on standardized tests (Woessmann, 2016). Correspondingly, the CCSS-induced learning loss in non-targeted subjects is equivalent to approximately 25 percent to 30 percent of a school year.

5.2 *Subgroup Analysis*

The negative effect of the CCSS on student achievement in non-targeted subjects need not be evenly distributed across subgroups. We hypothesize that students from underprivileged backgrounds may be disproportionately disadvantaged as it is more difficult for themselves, their parents, or their social environments in general to compensate for the reduction of instructional focus on non-targeted subjects.¹⁶ For example, parents from underprivileged backgrounds might be less able or might have less time to help their children with homework themselves or pay for private tuition. To test this hypothesis, we conduct subgroup analyses by students' demographic and socioeconomic characteristics.

As reported in Table 2, we find that the negative effect on student achievement in non-targeted subjects is not evenly distributed across subgroups. With respect to race/ethnicity, the student achievement of Hispanics and Blacks in non-targeted subjects is reduced disproportionately as a consequence of the CCSS, while there are almost no reform effects for Whites and Asians. Race and ethnicity aside, the negative effect is larger for students who qualify for subsidized lunch than for those who do not. Furthermore,

¹⁶In Section 7 on mechanisms, we use outcomes measuring instructional focus based on teacher survey data to show explicitly that the CCSS caused a reduction of instructional focus on non-targeted subjects.

students with English language learner status or disability status lose disproportionately. Taken together, these results indicate that students from groups typically regarded as socially, economically or physically underprivileged suffer most from CCSS exposure in terms of their achievement in non-targeted subjects.

We also perform subgroup analyses by subjects and grades. As far as subjects are concerned, the negative effect on student achievement in non-targeted subjects comes mostly from science, as we show in Table A.3. In terms of grades, the negative effect is mostly due to students from grade 4, see Table A.4. These two subgroup analyses should be interpreted with caution as testing frequencies and sample sizes are much larger for science (compared to civics, economics, geography, and history) for grades 4 and 8 (compared to grade 12). Still, the large subgroup effect for students in grade 4 makes intuitive sense as teachers in elementary school have the greatest flexibility in shifting the instructional focus.¹⁷

6 ROBUSTNESS

In this section, we test the robustness of the main result to account for four general types of concerns. First, we test the plausibility of the identifying assumption about parallel trends between treatment and control groups in the two-way fixed effects DD model in various ways in a series of econometric robustness tests. Second, we assess whether negative weighting induced by time-varying treatment effects affects our two-way fixed effects DD estimates and their interpretation. Third, we test whether re-defining our treatment variable to incorporate information about CCSS implementation changes our results. Fourth, we conduct a series of further specification checks to ensure that our

¹⁷Yet another subgroup analysis we perform is motivated by the hypothesis that states with high average student achievement levels before the adoption of CCSS might suffer more from CCSS adoption in terms of student achievement in non-targeted subjects relative to states that had been low-performing to begin with. To test this hypothesis, we perform subgroup analysis by quartiles of the states' pre-CCSS student achievement level. As presented in Table A.5, there is no linear effect pattern across quartiles. Although the effect difference between the lowest quartile and the highest quartile is negative which could support our hypothesis, we do not emphasize this subgroup finding as it does not hold in robustness checks on the econometric and treatment specifications as described in Section 6.

results do not hinge on arguably arbitrary specification choices in the main regression.

6.1 Robustness Tests on Parallel Trends Assumption

A first plausibility test for the identifying assumption of parallel trends in outcomes between treatment and control groups are event-study specifications. Here, the adoption of the CCSS in a given state and year is defined as an event. As depicted in Figure 2, no statistically significant pre-trend in student achievement in non-targeted subjects can be identified prior to the adoption of the CCSS. If at all, student achievement in these subjects was improving before the reform. In contrast to this insignificant positive pre-trend, student achievement declined substantially after the reform, both in terms of size and significance. This finding supports the validity of the parallel trends assumption and confirms the negative effect of the CCSS on student achievement in non-targeted subjects reported in the main analysis.

To further account for the possibility of state-specific trends and simultaneous shocks, Table 3 presents econometric robustness tests for the main regression results (shown again in column 1 to facilitate comparison) by first adding state-specific linear time trends (column 2), then adding state-specific quadratic time trends (column 3), and finally adding state-specific cubic time trends (column 4). In column 5, DDD estimates are reported in which a set of data on private school students is added to the sample. Here, all state-specific time trends are replaced with a fixed effect for school type (public school vs. private school), as well as with school-type-by-state, school-type-by-year, and state-by-year fixed effects. In column 6, we run the main two-way fixed effects DD model with a pooled sample of all students from public and private schools to get an estimate of the overall effect including potential spillovers.

We find that the negative and significant main effect is robust across specifications and becomes even slightly more negative in the models with state-specific time trends of various orders. The same is true for the DDD model. The model which is estimated based on the sample of both public and private school students yields an effect that is

slightly smaller than that of the main model, but still negative and significant. This series of econometric robustness checks demonstrates that the main effects are not affected by underlying state-specific trends in the data, endogenous selection between public and private school students, and shocks that occur simultaneously to the adoption of the CCSS as long as they affect both public and private school students equally.

While a large variety of state-specific shocks, including policy reforms in many areas, plausibly affect public and private schools students equally, the parallel trends assumption could still be violated by state-specific policies for public and private schooling. To test for this possibility, we perform robustness checks which explicitly control for state-specific time-varying policies for public schooling (Table A.6) and private schooling (Table A.7), respectively. The results are robust throughout. The first, second, and third specification in Table A.6 controls for NCLB waivers, state proficiency standards, and the adoption of the Next Generation Science Standards, respectively. The robustness of the CCSS exposure coefficient supports aforementioned interpretation that it is indeed the CCSS itself that drives the reported effects, and not waivers, other proficiency reforms or the NGSS. In addition, policies on homeschooling and compulsory schooling could indirectly affect public and private schooling decisions. To account for this possibility, we also present robustness checks in which we add controls for homeschooling and compulsory schooling. As can be seen in Table A.8 for homeschooling and Table A.9 for compulsory schooling, the results are robust. Taken together, the findings in this subsection indicate that neither underlying trends in the data nor shocks occurring simultaneously to the adoption of the CCSS give rise to our main result.

6.2 Robustness Tests on Time-Varying Treatment Effects

Another potential threat to identification of the presented two-way fixed effects DD models are time-varying treatment effects. The main estimate is a weighted sum of the average treatment effects in each state and year (i.e. of each 2x2 comparison), with weights that may be negative. These negative weights can cause the main regression coefficient to be

negative although all the average treatment effects are positive. Weights can be negative if already-treated units act as controls for later-treated units, in settings with time-varying treatment effects and staggered reform adoption.

Our event-study graphs presented above provide first insight into the potential bias induced by time-varying treatment effects, as they allow us to separate instantaneous from gradual reform effects. Event-studies are immune to bias from time-varying treatment effects as long as the pattern of effects is the same for all treatment cohorts. To explicitly explore the issue of time-varying treatment effects, we create a sample in which already-treated students never act as controls. Creating this sample is relatively straightforward in our setting as most states adopted the CCSS in 2010 or did not adopt the CCSS at all. By excluding the six states which adopted the CCSS in 2011 and 2012 from the sample, we transform our staggered setting into a non-staggered setting that is immune to negative weights. As shown in column 1 of Table A.10, the negative significant effect of the CCSS on student achievement in non-targeted subjects remains in this modified sample. This finding demonstrates that the main result is not driven by time-varying treatment effects and negative weights.

6.3 Robustness Tests on Treatment Definition

A different type of concern is that CCSS adoption and CCSS implementation could diverge. In our preferred treatment coding, we count all years as causing CCSS exposure for a student in a given state, in which the state had permanently adopted the CCSS before that year or at most in the same year. However, states that have adopted the CCSS permanently may not have implemented the CCSS comprehensively and thus may not be creating actual exposure. Conversely, states that have not adopted the CCSS permanently may have adopted and/or implemented the CCSS temporarily or partially.

To test whether our results hold if we define treatment based on CCSS implementation, we re-run our main regression using five different treatment variables, each capturing different information about the implementation of the CCSS. Under these treatment

definitions, a school year in a given state is defined as a school year with CCSS exposure if that state (i) expects teachers to fully incorporate the CCSS in their classroom instruction, (ii) followed at least two out of three CCSS implementation strategies (professional development, new instructional materials, joined testing consortium), (iii) observed an effective change in state standard content due to the adoption of the CCSS, which we define to mean that no state standard existed that closely resembles the CCSS before the adoption of the CCSS, (iv) adopted and/or implemented the CCSS at least temporarily, or (v) mandated standardized tests aligned to the CCSS. Further information on each treatment definition, its construction, data sources, including a table containing state-specific coding information for all treatment definitions are provided in Appendix C.

We present the results for each treatment definition (including our main result to facilitate comparison) in Table 4. We find negative point estimates throughout, ranging from -0.088 to -0.035 units of a standard deviation for the overall sample. However, most of them are statistically insignificant. We also show effects for the subsample of students in grade 4 for which we have observed the largest subgroup effects in the main analysis. For this hand-picked subsample, we find statistically significant effects throughout (and point estimates ranging from -0.177 to -0.098 units of a standard deviation). Taken together, these findings suggest that results using treatment definitions based on CCSS implementation rather than CCSS adoption lead to the same overall conclusion as the main results in Section 5.

6.4 *Further Specification Checks*

Furthermore, we want to confirm that our results indeed driven by the CCSS, and not by political factors. To test for this explicitly, we control for the party of the state governor in power, for the party of the state legislative chamber, and for both of these factors simultaneously. The relevant point estimate of the CCSS exposure on student achievement in non-targeted subjects is remarkably robust, indicating that our result is not biased by the political leaning of the state, see columns 2,3, and 4 of Table A.10.

In addition, we want to assess whether our results are robust to a number of modifications of our main regression. As indicated before, we set missing values of controls to zero and add separate explanatory binary variables to account for these missing values in our main regressions. The shares of missing values for the student control variables are below 10 percent for all variables except for parental education. For the latter approximately 40 percent of the values are missing, which can be mostly explained by the fact that this question was not asked in grade 4. To test whether the parental education control and its imputation affect the results, we run our main regression without controlling for parental education. As shown in column 5 of Table A.10, the effects do not differ meaningfully. As an additional robustness check, we do not impute missing values of any control variables (in addition to leaving parental education out of the set of control variables). As can be seen in column 6, the results are robust. Furthermore, we show that our results are robust to excluding charter schools from the sample of public schools, or omitting population weights, respectively, see columns 7 and 8.

7 MECHANISMS

To study what gave rise to the observed effect on student achievement in non-targeted subjects, we examine what changed in students' classrooms in these subjects due to the CCSS. To this end, we draw on teacher survey data, provided by the NAEP for a subset of waves and classrooms. This data is suitable for our analysis for several reasons. First, it contains a rich set of subject-specific questions on instructional focus in the classroom comprising instruction time, instructional resources, five measures of differentiated instruction, and four measures of the quality of teacher-student interactions. We note that the instructional focus outcomes could be endogenous to the reform and hence should be interpreted as changes in teachers' perceptions of classroom instruction rather than evidence based on administrative data (which is not available for these outcomes at the subject-state-year-level).

Second, the NAEP includes teacher background characteristics which we can use as

control variables and for subgroup analyses. Third, the NAEP teacher surveys are linked to the NAEP student achievement tests and student surveys. This link allows us to examine how instructional practices changed in non-targeted subjects according to the teachers who taught precisely the tested students from our main analysis. Fourth, the teacher surveys are standardized in the same way as the student surveys and achievement tests, making them comparable across states and years and thus suitable for a two-way fixed effects difference-in-differences approach. In fact, we can keep the empirical framework from the previous sections largely unchanged, but we use instructional focus outcomes instead of student outcomes and add teacher controls, thus ensuring methodological consistency with the previous sections.

Table 5 presents the results of CCSS exposure for instructional focus in non-targeted subjects. Overall, we find that the CCSS caused a reduction in instructional focus on the non-targeted subjects. Specifically, we observe negative significant effects of the CCSS on weekly instruction time, provision of instructional materials and resources, and two dimensions of the quality of teacher-student interactions. These two dimensions are setting and discussing goals with students. The extent of differentiated instruction did not change meaningfully. To illustrate the interpretation of the reported point estimates, we note that teachers of students who are fully exposed to the CCSS are 17 percentage points less likely to teach these students more than five hours per week in non-targeted subjects than teachers of students with no CCSS exposure, conditional on teacher characteristics, student characteristics as well as state, year, grade and subject fixed effects.¹⁸

This finding implies that the adoption of the CCSS has shifted the instructional focus away from the non-targeted subjects. This finding is in line with the results from Section 5, which show a decline in student achievement in these subjects. It is also consistent with previous literature showing that instructional inputs affect student achievement. Increases in instruction time (Taylor, 2014), instructional resources (Holden, 2016), and the quality

¹⁸The answer categories of the instructional outcome variables were coded differently in different survey waves of the NAEP. To ensure consistency across waves, we code the variables as reported in the footnote of Table 5.

of teacher-student interactions (Allen et al., 2011) have all been shown to positively affect student achievement. These instructional inputs can also interact. For example, the effect of instruction time on student achievement depends on student-teacher interactions (Rivkin and Schiman, 2015).

8 CONCLUSION

The recent empirical literature on school curricula has demonstrated that reforms of education standards can be a powerful policy tool to influence and modernize societies in the intended direction. However, as exemplified by this paper, such reforms can also have negative spillovers on non-targeted outcomes: Using the example of the CCSS adoption in math and ELA, we show that education standards can reduce student achievement in non-targeted subjects. With respect to mechanisms, we find that the negative spillover on student achievement in science and social studies was accompanied by a reduction of instructional time and focus on these subjects. These results can be interpreted as a warning for policy makers that topic-, subject-, and group-specific reforms of school curricula may come at substantial opportunity costs. Hence, we call for a careful evaluation of potential spillovers on non-targeted outcomes when designing targeted school curricula reforms. These insights may also translate to targeted education reforms more generally.

More specifically, we speculate that the negative spillover of the CCSS on student achievement in the non-targeted subjects could have been avoided if these subjects had been included in the CCSS. However, adopting such a comprehensive standard requires that the participating states agree on the educational content in these subjects. Political polarization would need to be overcome, as exemplified by controversies around how education standards should treat evolution theory, climate change, racism, and LGBTQ topics (Lerner, 2000; Cohen, 2020).

References

- Achieve Inc. (2013). “Closing the Expectations Gap: 2013 Annual Report on the Alignment of State K-12 Policies and Practice with the Demands of College and Careers”. *ERIC Clearinghouse*.
- Allen, Joseph P., Robert C. Pianta, Anne Gregory, Amori Yee Mikami, and Janetta Lun (2011). “An Interaction-Based Approach to Enhancing Secondary School Instruction and Student Achievement”. *Science* 333(6045), pp. 1034–1037.
- Allensworth, Elaine, Sarah Cashdollar, and Julia Gwynne (2021). “Improvements in Math Instruction and Student Achievement Through Professional Learning Around the Common Core State Standards in Chicago”. *AERA Open* 7, pp. 1–19.
- Arold, Benjamin W. (2022). “Evolution vs. Creationism in the Classroom: The Lasting Effects of Science Education”. *ifo Working Paper No. 379*.
- Barro, Robert J. (1991). “Economic Growth in a Cross Section of Countries”. *Quarterly Journal of Economics* 106(2), pp. 407–443.
- Barro, Robert J. and Jong Wha Lee (2013). “A New Data Set of Educational Attainment in the World, 1950–2010”. *Journal of Development Economics* 104, pp. 184–198.
- Barrow, Lisa, Lauren Sartain, and Marisa de la Torre (2020). “Increasing Access to Selective High Schools through Place-Based Affirmative Action: Unintended Consequences”. *American Economic Journal: Applied Economics* 12(4), pp. 135–63.
- Bay-Williams, Jennifer (2016). “Common Core Math in the K-8 Classroom: Results from a National Teacher Survey”. *Thomas B. Fordham Institute*.
- Bazzi, Samuel, Masyhur Hilmy, and Benjamin Marx (2020). “Religion, Education, and Development”. *National Bureau of Economic Research Working Paper Series No. 27073*.
- Bernheim, B.Douglas, Daniel M. Garrett, and Dean M. Maki (2001). “Education and saving:: The long-term effects of high school financial curriculum mandates”. *Journal of Public Economics* 80(3), pp. 435–465.

- Betts, Julian R. (1998). “The Impact of Educational Standards on the Level and Distribution of Earnings”. *American Economic Review* 88(1), pp. 266–275.
- Bishop, John Hillman (1997). “The Effect of National Standards and Curriculum-Based Exams on Achievement”. *American Economic Review* 87(2), pp. 260–264.
- Blazar, David, Blake Heller, Thomas J Kane, Morgan S. Polikoff, Douglas Staiger, Scott Carrell, Dan Goldhaber, Douglas Harris, Rachel Hitch, Kristian L Holden, et al. (2019). “Learning by the book: Comparing math achievement growth by textbook in six Common Core states”. *Center for Education Policy Research, Harvard University*.
- Bleiberg, Joshua (2021). “Does the Common Core Have a Common Effect? An Exploration of Effects on Academically Vulnerable Students”. *AERA Open* 7, pp. 1–18.
- Borghans, Lex, Bart H. H. Golsteyn, James J. Heckman, and John Eric Humphries (2016). “What Grades and Achievement Tests Measure”. *Proceedings of the National Academy of Sciences* 113(47), pp. 13354–13359.
- Borusyak, Kirill, Xavier Jaravel, and Jan Spiess (2021). “Revisiting Event Study Designs: Robust and Efficient Estimation”. *Working Paper*.
- Callaway, Brantly and Pedro H.C. Sant’Anna (2021). “Difference-in-Differences with multiple time periods”. *Journal of Econometrics* 225(2), pp. 200–230.
- Cameron, Lisa, Jennifer Seager, and Manisha Shah (2020). “Crimes Against Morality: Unintended Consequences of Criminalizing Sex Work”. *Quarterly Journal of Economics* 136(1), pp. 427–469.
- Cantoni, Davide, Yuyu Chen, David Y. Yang, Noam Yuchtman, and Y. Jane Zhang (2017). “Curriculum and Ideology”. *Journal of Political Economy* 125(2), pp. 338–392.
- Cantoni, Davide and Noam Yuchtman (2013). “The Political Economy of Educational Content and Development: Lessons from History”. *Journal of Development Economics* 104, pp. 233–244.
- Carmichael, Sheila Byrd, Gabrielle Martino, Kathleen Porter-Magee, and W Stephen Wilson (2010). “The State of State Standards-and the Common Core-in 2010”. *Thomas B. Fordham Institute*.

- CCSSI (2023). *Common Core State Standards Initiative: Standards in Your State*. Accessed on 2023-01-05 at: <http://www.corestandards.org/standards-in-your-state>.
- Chaisemartin, Clément de and Xavier D’Haultfœuille (2020). “Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects”. *American Economic Review* 110(9), pp. 2964–96.
- Clots-Figueras, Irma and Paolo Masella (2013). “Education, Language and Identity”. *Economic Journal* 123(570), pp. 332–357.
- Cohen, Eliot A (2020). “History, Critical and Patriotic: Americans Need a History That Educates but Also Inspires”. *Education Next* 20(2), pp. 8–18.
- Conger, Dylan, Alec I. Kennedy, Mark C. Long, and Raymond McGhee (2021). “The Effect of Advanced Placement Science on Students’ Skills, Confidence, and Stress”. *Journal of Human Resources* 56(1), pp. 93–124.
- Cortes, Kalena and Joshua Goodman (2014). “Ability-Tracking, Instructional Time, and Better Pedagogy: The Effect of Double-Dose Algebra on Student Achievement”. *American Economic Review* 104(5), pp. 400–405.
- Costrell, Robert M. (1994). “A Simple Model of Educational Standards”. *American Economic Review* 84(4), pp. 956–971.
- (1997). “Can Centralized Educational Standards Raise Welfare?” *Journal of Public Economics* 65(3), pp. 271–293.
- Daro, P., G. B. Hughes, and F. Stancavage (2015). “Study of the Alignment of the 2015 NAEP Mathematics Items at Grades 4 and 8 to the Common Core State Standards (CCSS) for Mathematics”. *Washington, D.C.: American Institutes for Research*.
- Dee, Thomas S. and Brian Jacob (2011). “The impact of no Child Left Behind on student achievement”. *Journal of Policy Analysis and Management* 30(3), pp. 418–446.
- Fouka, Vasiliki (2019). “Backlash: The Unintended Effects of Language Prohibition in U.S. Schools after World War I”. *Review of Economic Studies* 87(1), pp. 204–239.
- Fuchs-Schündeln, Nicola and Paulo Masella (2016). “Long-Lasting Effects of Socialist Education”. *Review of Economics and Statistics* 98(3), pp. 428–441.

- Gao, Niu and Julien Lafortune (2019). “Common Core State Standards in California: Evaluating Local Implementation and Student Outcomes.” *Public Policy Institute of California*.
- Goodman, Joshua (2019). “The Labor of Division: Returns to Compulsory High School Math Coursework”. *Journal of Labor Economics* 37(4), pp. 1141–1182.
- Goodman-Bacon, Andrew (2021). “Difference-in-differences with variation in treatment timing”. *Journal of Econometrics* 225(2). Themed Issue: Treatment Effect 1, pp. 254–277.
- Hamlin, Daniel and Paul E Peterson (2018). “Have states maintained high expectations for student performance? An analysis of 2017 state proficiency standards”. *Education Next* 18(4), pp. 42–49.
- Hanushek, Eric A, Paul E Peterson, and Ludger Woessmann (2012). “Achievement Growth: International and US State Trends in Student Performance”. *Program on Education Policy and Government Report No.: 12-03, Taubman Center for State and Local Government, Harvard Kennedy School*.
- Hanushek, Eric A. and Ludger Woessmann (2008). “The Role of Cognitive Skills in Economic Development”. *Journal of Economic Literature* 46(3), pp. 607–668.
- (2012). “Do Better Schools Lead to More Growth? Cognitive Skills, Economic Outcomes, and Causation”. *Journal of Economic Growth* 17(4), pp. 267–321.
- Heckman, James J. and Tim Kautz (2012). “Hard evidence on soft skills”. *Labour Economics* 19(4), pp. 451–464.
- Holden, Kristian L. (2016). “Buy the Book? Evidence on the Effect of Textbook Funding on School-Level Achievement”. *American Economic Journal: Applied Economics* 8(4), pp. 100–127.
- Jackson, C. Kirabo (2018). “What Do Test Scores Miss? The Importance of Teacher Effects on Non-Test Score Outcomes”. *Journal of Political Economy* 126(5), pp. 2072–2107.

- Jackson, C. Kirabo, Cora Wigger, and Heyu Xiong (2021). “Do School Spending Cuts Matter? Evidence from the Great Recession”. *American Economic Journal: Economic Policy* 13(2), pp. 304–35.
- Jacob, Brian A. (2005). “Accountability, incentives and behavior: the impact of high-stakes testing in the Chicago Public Schools”. *Journal of Public Economics* 89(5), pp. 761–796.
- Jerald, Craig D (2008). “Benchmarking for Success: Ensuring US Students Receive a World-Class Education”. *Washington, DC: National Governors Association*.
- Johnson, Carla C (2007). “Effective science teaching, professional development and no child left behind: Barriers, dilemmas, and reality”. *Journal of Science Teacher Education* 18(2), pp. 133–136.
- Jordan, Marty P and Matt Grossmann (2020). “The Correlates of State Policy Project v.2.2.” *East Lansing, MI: Institute for Public Policy and Social Research (IPPSR)*.
- Kane, Thomas J, Antoniya M Owens, William H Marinell, Daniel RC Thal, and Douglas O Staiger (2016). “Teaching higher: Educators’ perspectives on Common Core implementation”. *Cambridge, MA: Center for Education Policy Research*, pp. 9–15.
- Lafortune, Julien, Jesse Rothstein, and Diane Whitmore Schanzenbach (2018). “School finance reform and the distribution of student achievement”. *American Economic Journal: Applied Economics* 10(2), pp. 1–26.
- Lerner, Lawrence S (2000). “Good and Bad Science in US Schools”. *Nature* 407(6802), pp. 287–290.
- Lott, John R. Jr. (1999). “Public Schooling, Indoctrination, and Totalitarianism”. *Journal of Political Economy* 107(6), pp. 127–157.
- Loveless, Tom (2014). “The Brown Center Report on Education. How well are American students learning. Part III: A progress report on the Common Core”. *The Brookings Institution*.
- (2015). “The Brown Center Report on Education. How well are American students learning? Part II: Measuring effects of the Common Core”. *The Brookings Institution*.

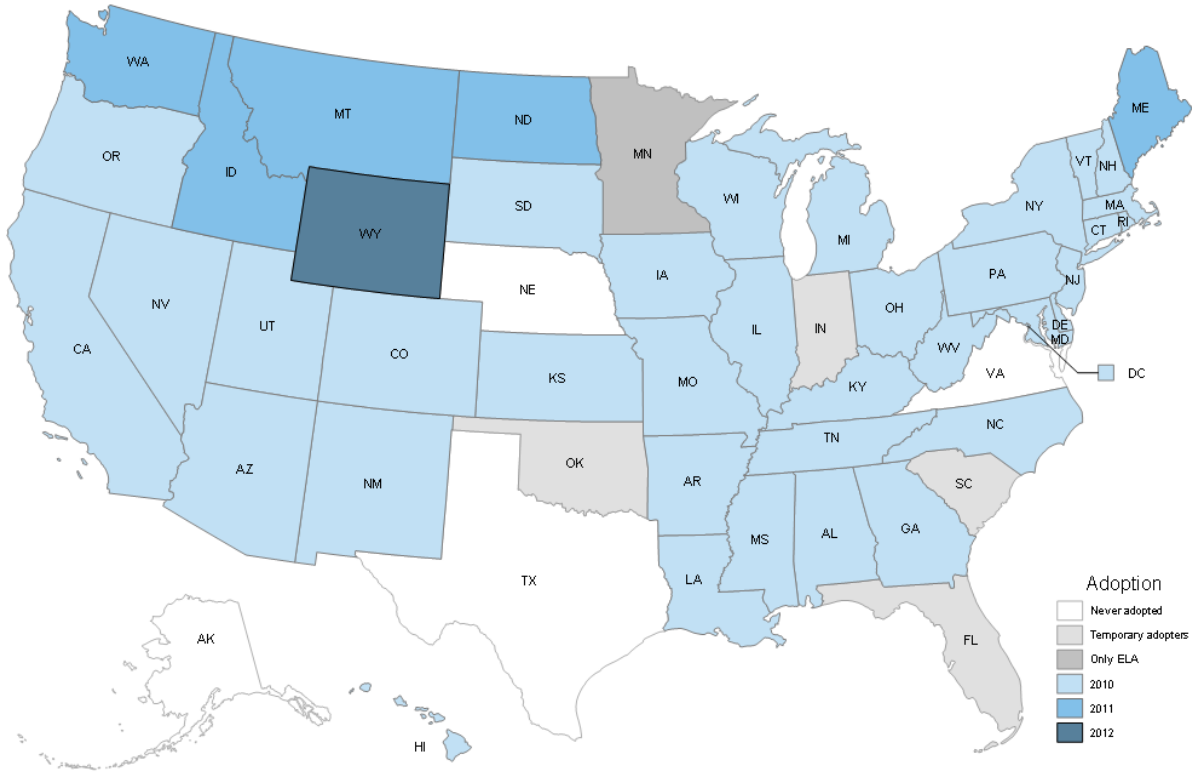
- Loveless, Tom (2016). “The Brown Center Report on Education. How well are American students learning? Part I: Reading and Math in the Common Core Era”. *The Brookings Institution*.
- Misco, Thomas (2005). “In response to NCLB: A case for retaining the social studies”. *Essays in Education* 15(1), p. 2.
- National Council for Social Studies (2008). “Curriculum Guidelines for Social Studies Teaching and Learning”. *National Council for Social Studies Position Statement*.
- Neal, Derek and Diane Whitmore Schanzenbach (2010). “Left Behind by Design: Proficiency Counts and Test-Based Accountability”. *Review of Economics and Statistics* 92(2), pp. 263–283.
- Opfer, V Darleen, Julia H Kaufman, and Lindsey E Thompson (2016). “Implementation of K–12 state standards for mathematics and English language arts and literacy”. *Santa Monica, CA: RAND*.
- Polikoff, Morgan, Elaine Lin Wang, Shira Korn Haderlein, Julia H Kaufman, Ashley Woo, Daniel Silver, and V Darleen Opfer (2020). “Exploring Coherence in English Language Arts Instructional Systems in the Common Core Era”. *RAND Corporation Research Report. RR-A279-1*.
- Polikoff, Morgan S. (2015). “How Well Aligned Are Textbooks to the Common Core Standards in Mathematics?” *American Educational Research Journal* 52(6), pp. 1185–1211.
- (2017). “Is Common Core “Working”? And Where Does Common Core Research Go From Here?” *AERA Open* 3(1), pp. 1–6.
- Polleck, Jody N. and Jill V. Jeffery (2017). “Common Core Standards and their Impact on Standardized Test Design: A New York Case Study”. *High School Journal* 101(1), pp. 1–26.
- Reardon, S. F., A. D. Ho, B. R. Shear, E. M. Fahle, D. Kalogrides, H. Jang, and B. Chavez (2021). “Stanford Education Data Archive (Version 4.0)”.

- Reback, Randall (2008). “Teaching to the rating: School accountability and the distribution of student achievement”. *Journal of Public Economics* 92(5), pp. 1394–1415.
- Reback, Randall, Jonah Rockoff, and Heather L. Schwartz (2014). “Under Pressure: Job Security, Resource Allocation, and Productivity in Schools under No Child Left Behind”. *American Economic Journal: Economic Policy* 6(3), pp. 207–41.
- Rivkin, Steven G. and Jeffrey C. Schiman (2015). “Instruction Time, Classroom Quality, and Academic Achievement”. *Economic Journal* 125(588), pp. 425–448.
- Romer, Paul M. (1990). “Human Capital and Growth: Theory and Evidence”. *Carnegie-Rochester Conference Series on Public Policy* 32, pp. 251–286.
- Ross, E, N Gerber, H Jarmolowski, K Lakis, N Ledyard, L Staresina, and C Worth (2017). “State teacher policy yearbook”. *Washington, DC: National Council on Teacher Quality*.
- Rozo, Sandra V. (2020). “Unintended effects of illegal economic activities: Illegal gold mining and malaria”. *World Development* 136, pp. 105–119.
- Schmidt, William H. and Richard T. Houang (2012). “Curricular Coherence and the Common Core State Standards for Mathematics”. *Educational Researcher* 41(8), pp. 294–308.
- Scholastic (2014). “Teachers’ views on the Common Core State Standards one year later”. *The Bill & Melinda Gates Foundation*.
- Settlage, John and Lee Meadows (2002). “Standards-based reform and its unintended consequences: Implications for science education within America’s urban schools”. *Journal of Research in Science Teaching* 39(2), pp. 114–127.
- Shakeel, M Danish and Paul E Peterson (2022). “A half century of progress in US student achievement: Agency and Flynn effects, ethnic and SES differences”. *Educational Psychology Review* 34(3), pp. 1255–1342.
- Song, Mengli, Rui Yang, and Michael Garet (2019). “Effects of States’ Implementation of College-and Career-Ready Standards on Student Achievement”. Working Paper Series.

- Sorens, Jason, Fait Muedini, and William P Ruger (2008). “US state and local public policies in 2006: A new database”. *State Politics & Policy Quarterly* 8(3), pp. 309–326.
- Sun, Liyang and Sarah Abraham (2021). “Estimating dynamic treatment effects in event studies with heterogeneous treatment effects”. *Journal of Econometrics* 225(2), pp. 175–199.
- Taylor, Eric (2014). “Spending more of the school day in math class: Evidence from a regression discontinuity in middle school”. *Journal of Public Economics* 117, pp. 162–181.
- Webber, Ann, Patricia Troppe, Anthony Milanowski, Babette Gutmann, Elizabeth Reisner, and Margaret Goertz (2014). “State Implementation of Reforms Promoted under the Recovery Act. A Report from Charting the Progress of Education Reform: An Evaluation of the Recovery Act’s Role”. *National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U. S. Department of Education*.
- Winters, Marcus A., Julie R. Trivitt, and Jay P. Greene (2010). “The impact of high-stakes testing on student proficiency in low-stakes subjects: Evidence from Florida’s elementary science exam”. *Economics of Education Review* 29(1), pp. 138–146.
- Woessmann, Ludger (2016). “The Importance of School Systems: Evidence from International Differences in Student Achievement”. *Journal of Economic Perspectives* 30(3), pp. 3–32.
- Xu, Zeyu and Kennan Ceba (2018). “Getting College-Ready during State Transition toward the Common Core State Standards”. *Teachers College Record* 120(6), pp. 1–36.

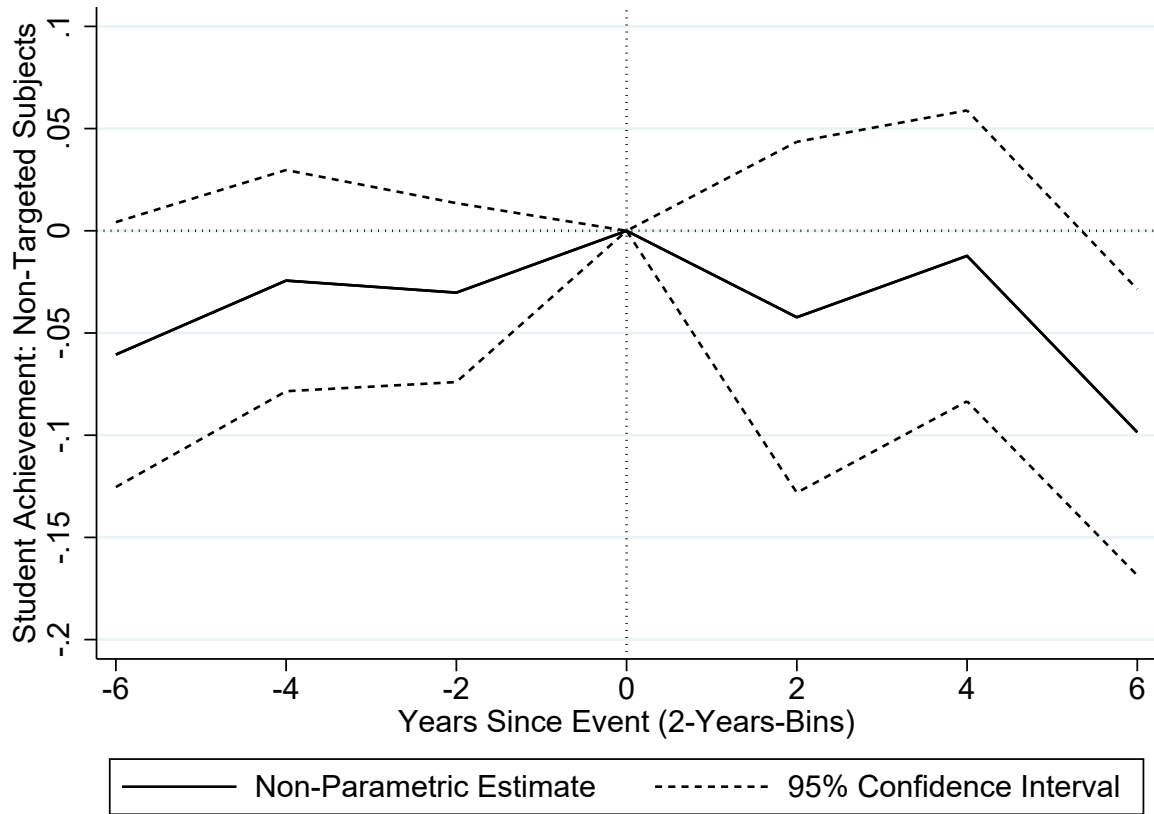
MAIN FIGURES AND TABLES

Figure 1 – CCSS adoption map



Note: Map depicts state-level adoption of CCSS. Data sources: Achieve Inc. (2013), Bleiberg (2021), and CCSSI (2023)

Figure 2 – Event-study graph: Non-targeted subjects



Note: Coefficients from non-parametric event-study regressions and their 95% confidence intervals. Dependent variable: Standardized student achievement in subjects not targeted by the CCSS (Pool of science, civics, economics, geography, and history). Controls: Indicator variables for gender, races/ethnicities, English language learner status, disability status, subsidized lunch status, parental education, home possessions (separate indicator variables for computer and books) as well as state, test year, grade and subject fixed effects. Regressions use population weights and standard errors clustered at the state level. Numbers on horizontal axis refer to respective two-year bins; i.e. 2 = first two years of treatment (year 0 = excluded category). The p values of omnibus hypothesis tests of zero pre- and post-event effects are 0.312 and 0.001, respectively. Data source: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress

Table 1 – Effect of CCSS exposure on student achievement in non-targeted subjects

	(1)	(2)	(3)
CCSS Exposure	0.105*** (0.036)	0.117*** (0.043)	-0.079** (0.036)
State and Year FEs	NO	NO	YES
Controls	NO	YES	YES
Adj. R-squared	0.001	0.379	0.390
Observations	1,103,630	1,103,630	1,103,630

Note: Each entry is from a separate two-way fixed effects regression model. Dependent variable: Standardized student achievement in subjects not targeted by the CCSS (Pool of science, civics, economics, geography, and history). Explanatory variable: Share of schooling years a student was exposed to CCSS (at the time of testing). Controls: Indicator variables for gender, races/ethnicities, English language learner status, disability status, subsidized lunch status, parental education, home possessions (separate indicator variables for computer and books) as well as grade and subject fixed effects. Regressions use population weights and standard errors clustered at the state level. Single, double, and triple asterisks indicate statistical significance at the 10%, 5%, and 1% levels, respectively. Data source: See Figure 2

Table 2 – Effect of CCSS exposure on student achievement in non-targeted subjects, subgroups by student characteristics

	Gender		Race/Ethnicity				Subsidized Lunch Status		English Language Learner Status		Disability Status		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Female		Male	White	Black	Hispanic	Asian	Other	Yes	No	Yes	No	Yes	No
CCSS Exposure	-0.079** (0.037)	-0.080** (0.036)	-0.016 (0.039)	-0.111*** (0.037)	-0.181*** (0.041)	-0.012 (0.061)	0.016 (0.071)	-0.096*** (0.031)	-0.041 (0.044)	-0.201*** (0.073)	-0.065* (0.035)	-0.176*** (0.049)	-0.068* (0.037)
State and Year FEs	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Adj. R-squared	0.395	0.390	0.272	0.241	0.331	0.392	0.362	0.303	0.301	0.182	0.348	0.308	0.353
Observations	544,410	559,210	631,640	184,680	194,250	55,130	37,930	513,910	589,710	69,190	1,034,440	128,590	975,030

Note: Each entry is from a separate two-way fixed effects regression model. Dependent variable: Standardized student achievement in subjects not targeted by the CCSS (Pool of science, civics, economics, geography, and history). Explanatory variable: Share of schooling years a student was exposed to CCSS (at the time of testing). Controls: Indicator variables for gender, races/ethnicities, English language learner status, disability status, subsidized lunch status, parental education, home possessions (separate indicator variables for computer and books) as well as grade and subject fixed effects. Regressions use population weights and standard errors clustered at the state level. Single, double, and triple asterisks indicate statistical significance at the 10%, 5%, and 1% levels, respectively. Data source: See Figure 2

Table 3 – Effect of CCSS exposure on student achievement in non-targeted subjects, econometric robustness

	(1)	(2)	(3)	(4)	(5)	(6)
	DD	(1) + linear state trends	(2) + quadratic state trends	(3) + cubic state trends	DDD	DD with private schools
CCSS Exposure	-0.079** (0.036)	-0.117** (0.044)	-0.100** (0.047)	-0.095** (0.045)	-0.090** (0.044)	-0.074** (0.033)
State and Year FEs	YES	YES	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES	YES
Adj. R-squared	0.390	0.391	0.392	0.393	0.394	0.390
Observations	1,103,630	1,103,630	1,103,630	1,103,630	1,135,960	1,135,960

Note: Each entry is from a separate two-way fixed effects regression model, where Model (1) is the baseline model, Models (2), (3), (4) subsequently add linear, quadratic, and cubic state-specific time trends, and Model (5) presents a triple-difference model where school type (public vs. private school students), school type*state, school type*year, and state*year fixed effects replace all state-specific time trends. Model (6) estimates the basic two-way fixed effects model on a sample of public and private school students. Dependent variable: Standardized student achievement in subjects not targeted by the CCSS (Pool of science, civics, economics, geography, and history). Explanatory variable: Share of schooling years a student was exposed to CCSS (at the time of testing). Controls: Indicator variables for gender, races/ethnicities, English language learner status, disability status, subsidized lunch status, parental education, home possessions (separate indicator variables for computer and books) as well as grade and subject fixed effects. Regressions use population weights and standard errors clustered at the state level. Single, double, and triple asterisks indicate statistical significance at the 10%, 5%, and 1% levels, respectively. Data source: See Figure 2

Table 4 – Effect of CCSS exposure on student achievement in non-targeted subjects, robustness using different definitions of treatment implementation

	CCSS adoption		CCSS implementation requirement		CCSS implementation strategies		Effective CCSS implementation		Include temporary CCSS adopters and implementers		CCSS-aligned testing	
	(1) All grades	(2) Only grade 4	(3) All grades	(4) Only grade 4	(5) All grades	(6) Only grade 4	(7) All grades	(8) Only grade 4	(9) All grades	(10) Only grade 4	(11) All grades	(12) Only grade 4
CCSS Exposure	-0.070** (0.036)	-0.134*** (0.036)	-0.088 (0.079)	-0.177** (0.076)	-0.042 (0.041)	-0.098* (0.050)	-0.075** (0.035)	-0.107*** (0.039)	-0.035 (0.045)	-0.111** (0.055)	-0.085 (0.057)	-0.118** (0.054)
State and Year FEs	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Adj. R-squared	0.390	0.376	0.390	0.376	0.390	0.376	0.390	0.376	0.390	0.376	0.390	0.376
Observations	1,103,630	434,440	1,103,630	434,440	1,103,630	434,440	1,103,630	434,440	1,103,630	434,440	1,103,630	434,440

Note: Each entry is from a separate two-way fixed effects regression model. Dependent variable: Standardized student achievement in subjects not targeted by the CCSS (Pool of science, civics, economics, geography, and history). Explanatory variables: Share of schooling years a student was exposed to CCSS (at the time of testing), where in Models 1 and 2 (CCSS adoption, baseline model) each schooling year counts as exposed in a given state in which the state adopted the CCSS permanently before that year or in the same year according to Achieve Inc. (2013) and CCSSI (2023), where Models 3 and 4 (CCSS implementation requirement) each schooling year counts as exposed in a given state in which the state expects teachers to fully incorporate CCSS into classroom instruction in grades K-12 in English language arts and mathematics according to Achieve Inc. (2013) and CCSSI (2023), where in Models 5 and 6 (CCSS implementation strategies) each schooling year counts as exposed in a given state if state education agency officials report that their state pursued at least two out of three CCSS implementation strategies (professional development, new instructional materials, joined testing consortium) as reported in Webber et al. (2014), where in Models 7 and 8 (Effective CCSS implementation) each schooling year counts as exposed in a given state in which the state implemented an effective change in state standard content through the adoption of CCSS which we define as not having had a state standard in place before the adoption of CCSS whose academic rigor is "too close to call" in comparison with CCSS (Carmichael et al., 2010) for the set of states adopting CCSS according to Achieve Inc. (2013) and CCSSI (2023), where in Models 9 and 10 (Include temporary CCSS adopters) each schooling year counts as exposed in a given state in which a state adopted and/or implemented CCSS at least temporarily according to Bleiberg (2021); and where in Models 11 and 12 (CCSS-aligned testing) each schooling year counts as exposed in a given state in which the state adopted CCSS-aligned standardized testing including field and transitional tests according to our own research (see Table C.14 for state-specific details of CCSS-aligned testing). Table C.15 provides state-specific coding information on all treatment definitions. Controls: Indicator variables for gender, races/ethnicities, English language learner status, disability status, subsidized lunch status, parental education, home possessions (separate indicator variables for computer and books) as well as grade and subject fixed effects. Regressions use population weights and standard errors clustered at the state level. Single, double, and triple asterisks indicate statistical significance at the 10%, 5%, and 1% levels, respectively. Data source: See Figure 2

Table 5 – Effect of CCSS exposure on instructional focus in non-targeted subjects

	Instructional Resources		Differentiated Instruction				Teacher-Student Interactions				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	Instruction Time	Instructional Resources	Standards	Material	Activities	Methods	Pace	Discuss Students' Performance	Set goals	Discuss goals	Adjust teaching
CCSS Exposure	-0.171*** (0.055)	-0.116*** (0.030)	-0.027* (0.015)	0.021 (0.016)	-0.022 (0.013)	-0.015 (0.022)	0.006 (0.016)	-0.017 (0.035)	-0.035* (0.019)	-0.047** (0.023)	-0.020 (0.023)
State and Year FEs	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Adj. R-squared	0.090	0.043	0.053	0.073	0.064	0.167	0.122	0.109	0.091	0.089	0.058
Observations	847,830	785,660	555,490	556,350	556,070	556,440	556,110	554,820	554,490	554,360	554,250

Note: Each entry is from a separate two-way fixed effects regression model. Dependent variables, by columns: Probability that students' teachers report about subject in question (Model 1) that students receive more than five hours of weekly instruction, (Model 2) that their school system provides them with all or most materials and other resources they need for the instruction, (Model 3) that they set differentiated standards for some students at least to a moderate extent, (Model 4) that they use differentiated materials for some students at least to a moderate extent, (Model 5) engage some students in differentiated activities at least to a moderate extent, (Model 6) that they use differentiated methods for some students at least to a moderate extent, (Model 7) change pace for some students at least to a moderate extent, (Model 8) that they discuss the student's current level of performance at least once a month, (Model 9) that they set goals for specific progress the student would like to make at least once a month, (Model 10) that they discuss progress the student has made toward goals previously set at least once a month, (Model 11) that they determine how to adjust their teaching strategies to meet the student's current learning needs and to reflect the student's future goals at least once a month. Explanatory variable: Share of schooling years teacher's students were exposed to CCSS (at the time of testing). Controls: Indicator variables for teacher race/ethnicity, teacher education, teacher certification (separate indicator variables for certification of National Board for Professional Teaching Standards, and for alternative certification), and teacher experience, as well as student controls for gender, races/ethnicities, English language learner status, disability status, subsidized lunch status, parental education, home possessions (separate indicator variables for computer and books) as well as grade and subject fixed effects. Regressions use population weights and standard errors clustered at the state level. Single, double, and triple asterisks indicate statistical significance at the 10%, 5%, and 1% levels, respectively. Data source: See Figure 2

A SUPPLEMENTARY FIGURES AND TABLES

Table A.1 – List of grades of NAEP tests for non-targeted subjects

Year	Non-targeted Subjects				
	Science	Civics	Economics	Geography	History
2005	4, 8				
2006		4, 8, 12	12		4, 8, 12
2007					
2008					
2009	4, 8, 12				
2010		4, 8, 12		4, 8, 12	4, 8, 12
2011	8				
2012			12		
2013					
2014		8		8	8
2015	4, 8, 12				

Note: NAEP student achievement data in non-targeted subjects at the subject-by-year-by-grade level. Data source: See Figure 2

Table A.2 – Descriptive statistics

	Mean	Std. Dev.	Min.	Max.
<i>Student Achievement Outcomes:</i>				
Science	0.08	1.03	-4.76	4.51
Civics	0.03	0.98	-4.49	3.16
Economics	0.02	0.97	-4.33	3.70
Geography	0.00	1.00	-5.15	3.90
History	0.04	0.98	-4.64	3.32
<i>Student Controls:</i>				
Female	0.49	0.50	0.00	1.00
Race/Ethnicity: White	0.57	0.49	0.00	1.00
Race/Ethnicity: Black	0.15	0.36	0.00	1.00
Race/Ethnicity: Hispanic	0.20	0.40	0.00	1.00
Race/Ethnicity: Asian	0.05	0.22	0.00	1.00
Race/Ethnicity: Other	0.03	0.16	0.00	1.00
English Language Learner	0.06	0.24	0.00	1.00
Disabled	0.11	0.31	0.00	1.00
Subsidized Lunch	0.43	0.50	0.00	1.00
Parental Education: Did not finish High School	0.08	0.28	0.00	1.00
Parental Education: Graduated High School	0.18	0.39	0.00	1.00
Parental Education: Some education after High School	0.20	0.40	0.00	1.00
Parental Education: Graduated College	0.53	0.50	0.00	1.00
Computer at Home	0.90	0.30	0.00	1.00
Books at Home: 0-10	0.13	0.34	0.00	1.00
Books at Home: 11-25	0.22	0.41	0.00	1.00
Books at Home: 26-100	0.35	0.48	0.00	1.00
Books at Home: >100	0.30	0.46	0.00	1.00

Note: Continuation on next page

Descriptive statistics (continued)

	Mean	Std. Dev.	Min.	Max.
<i>Instructional Focus Outcomes:</i>				
Instruction Time	0.41	0.49	0.00	1.00
Instructional Resources	0.60	0.49	0.00	1.00
Differentiated Instruction: Standards	0.44	0.50	0.00	1.00
Differentiated Instruction: Material	0.65	0.48	0.00	1.00
Differentiated Instruction: Activities	0.41	0.49	0.00	1.00
Differentiated Instruction: Methods	0.62	0.49	0.00	1.00
Differentiated Instruction: Pace	0.59	0.49	0.00	1.00
Teacher dedication: Discuss students' performance	0.56	0.50	0.00	1.00
Teacher dedication: Set goals	0.40	0.49	0.00	1.00
Teacher dedication: Discuss goals	0.41	0.49	0.00	1.00
Teacher dedication: Adjust teaching	0.64	0.48	0.00	1.00
<i>Teacher Controls:</i>				
Teacher Race/Ethnicity: White	0.83	0.38	0.00	1.00
Teacher Race/Ethnicity: Black	0.07	0.25	0.00	1.00
Teacher Race/Ethnicity: Hispanic	0.06	0.24	0.00	1.00
Teacher Race/Ethnicity: Asian	0.02	0.15	0.00	1.00
Teacher Race/Ethnicity: Other	0.01	0.11	0.00	1.00
Teacher Education: Bachelor or less	0.50	0.50	0.00	1.00
Teacher Education: Master or more	0.50	0.50	0.00	1.00
NBPTS Teacher Certificate: Yes	0.13	0.34	0.00	1.00
NBPTS Teacher Certificate: Working towards	0.02	0.15	0.00	1.00
NBPTS Teacher Certificate: No	0.85	0.36	0.00	1.00
Alternative Teacher Certificate: Yes	0.13	0.33	0.00	1.00
Alternative Teacher Certificate: No	0.87	0.33	0.00	1.00
Teacher Experience: 2 years or less	0.09	0.28	0.00	1.00
Teacher Experience: 3-5 years	0.14	0.34	0.00	1.00
Teacher Experience: 6-10 years	0.22	0.42	0.00	1.00
Teacher Experience: 11-20 years	0.27	0.44	0.00	1.00
Teacher Experience: 21 years or more	0.28	0.45	0.00	1.00

Note: Descriptive statistics (mean, standard deviation, minimum, maximum) for main treatment, outcome, and control variables. NBPTS stands for National Board for Professional Teaching Standards. Data source: See Figure 2

Table A.3 – Effect of CCSS exposure on student achievement in non-targeted subjects, subgroups by subjects

	(1)	(2)	(3)	(4)	(5)
	Science	Civics	Economics	Geography	History
CCSS Exposure	-0.096** (0.042)	0.010 (0.087)	-0.052 (0.307)	-0.036 (0.079)	-0.004 (0.097)
State and Year FEs	YES	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES
Adj. R-squared	0.402	0.391	0.381	0.429	0.380
Observations	931,600	55,150	19,930	32,130	64,810

Note: Each entry is from a separate two-way fixed effects regression model. Dependent variable: Standardized student achievement in subject indicated in the column header. Explanatory variable: Share of schooling years a student was exposed to CCSS (at the time of testing). Controls: Indicator variables for gender, races/ethnicities, English language learner status, disability status, subsidized lunch status, parental education, home possessions (separate indicator variables for computer and books) as well as grade fixed effects. Regressions use population weights and standard errors clustered at the state level. Single, double, and triple asterisks indicate statistical significance at the 10%, 5%, and 1% levels, respectively. Data source: See Figure 2

Table A.4 – Effect of CCSS exposure on student achievement in non-targeted subjects, subgroups by grades

	(1)	(2)	(3)
	Grade 4	Grade 8	Grade 12
CCSS Exposure	-0.134*** (0.036)	-0.007 (0.062)	-0.005 (0.072)
State and Year FEs	YES	YES	YES
Controls	YES	YES	YES
Adj. R-squared	0.376	0.433	0.370
Observations	434,440	582,590	86,600

Note: Each entry is from a separate two-way fixed effects regression model. Dependent variable: Standardized student achievement in subjects not targeted by the CCSS (Pool of science, civics, economics, geography, and history). Explanatory variable: Share of schooling years a student was exposed to CCSS (at the time of testing). Controls: Indicator variables for gender, races/ethnicities, English language learner status, disability status, subsidized lunch status, parental education, home possessions (separate indicator variables for computer and books) as well as subject fixed effects. Regressions use population weights and standard errors clustered at the state level. Single, double, and triple asterisks indicate statistical significance at the 10%, 5%, and 1% levels, respectively. Data source: See Figure 2

Table A.5 – Effect of CCSS exposure on student achievement in non-targeted subjects, subgroups by quartiles of states’ student achievement before 2010

	(1) Quartile 1 (lowest)	(2) Quartile 2	(3) Quartile 3	(4) Quartile 4 (highest)
CCSS Exposure	0.004 (0.046)	-0.097* (0.048)	-0.061 (0.051)	-0.041 (0.029)
State and Year FEs	YES	YES	YES	YES
Controls	YES	YES	YES	YES
Adj. R-squared	0.413	0.365	0.353	0.380
Observations	315,160	334,740	233,430	220,300

Note: Each entry is from a separate two-way fixed effects regression model. Sample of Quartile 1 subgroup includes students from states in the lowest quartile with respect to average student achievement in years before 2010. Sample of Quartile 2 subgroup includes students from states in the second lowest quartile with respect to average student achievement in years before 2010. Quartile 3 and 4 defined accordingly. Dependent variable: Standardized student achievement in subjects not targeted by the CCSS (Pool of science, civics, economics, geography, and history). Explanatory variable: Share of schooling years a student was exposed to CCSS (at the time of testing). Controls: Indicator variables for gender, races/ethnicities, English language learner status, disability status, subsidized lunch status, parental education, home possessions (separate indicator variables for computer and books) as well as grade and subject fixed effects. Regressions use population weights and standard errors clustered at the state level. Single, double, and triple asterisks indicate statistical significance at the 10%, 5%, and 1% levels, respectively. Data source: See Figure 2

Table A.6 – Effect of CCSS exposure on student achievement in non-targeted subjects, robustness with additional controls for public schooling policies

	Control for:							
	(1) NCLB/ESSA Waivers	(2) State Proficiency	(3) NGSS Adoption	(4) Expenditures	(5) Teacher Policies	(6) School Choice	(7) Evolution	(8) Charter Schools
CCSS Exposure	-0.074** (0.035)	-0.083** (0.037)	-0.066* (0.034)	-0.077** (0.035)	-0.098** (0.037)	-0.075** (0.037)	-0.079** (0.036)	-0.080** (0.036)
State and Year FEs	YES	YES	YES	YES	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES	YES	YES	YES
Adj. R-squared	0.390	0.390	0.390	0.391	0.397	0.390	0.390	0.390
Observations	1,103,630	1,089,080	1,103,630	1,103,620	769,410	1,099,880	1,099,880	1,099,880

Note: Each entry is from a separate two-way fixed effects regression model. Dependent variable: Standardized student achievement in subjects not targeted by the CCSS (Pool of science, civics, economics, geography, and history). Explanatory variable: Share of schooling years a student was exposed to CCSS (at the time of testing). Controls: Indicator variables for gender, races/ethnicities, English language learner status, disability status, subsidized lunch status, parental education, home possessions (separate indicator variables for computer and books) as well as grade and subject fixed effects. Additional policy controls (state-by-year level, unless otherwise stated): Model 1 controls for NCLB/ESSA requirements waiver; Model 2 controls for percentile ranking of state’s proficiency standards (average over math and ELA proficiency for grades 4 and 8; missing years are imputed as average of adjacent years); Model 3 controls for adoption of Next Generation Science Standards or of standards based on Next Generation Science Standards framework; Model 4 controls for district-by-year-level per-pupil education expenditures in logarithmized dollars; Model 5 controls for index of teacher quality policies; Model 6 controls for public school choice laws; Model 7 controls for laws permitting public school teachers to teach ‘weaknesses of evolution’; Model 8 controls for charter school laws. Regressions use population weights and standard errors clustered at the state level. Single, double, and triple asterisks indicate statistical significance at the 10%, 5%, and 1% levels, respectively. Data source: National Center for Education Statistics (Local Education Agency (School District) Finance Survey F-33); Sorens et al. (2008); Ross et al. (2017); Hamlin and Peterson (2018); Jordan and Grossmann (2020); See Figure 2

Table A.7 – Effect of CCSS exposure on student achievement in non-targeted subjects, robustness with additional controls for private schooling policies

	Control for:					
	(1) State Approval	(2) Licensure of Teachers	(3) Registration	(4) Curriculum	(5) Tax Credits	(6) Vouchers
CCSS Exposure	-0.079** (0.036)	-0.080** (0.036)	-0.075** (0.036)	-0.080** (0.036)	-0.088*** (0.028)	-0.064* (0.035)
State and Year FEs	YES	YES	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES	YES
Adj. R-squared	0.390	0.390	0.390	0.390	0.390	0.390
Observations	1,099,880	1,099,880	1,099,880	1,099,880	1,099,880	1,099,880

Note: Each entry is from a separate two-way fixed effects regression model. Dependent variable: Standardized student achievement in subjects not targeted by the CCSS (Pool of science, civics, economics, geography, and history). Explanatory variable: Share of schooling years a student was exposed to CCSS (at the time of testing). Controls: Indicator variables for gender, races/ethnicities, English language learner status, disability status, subsidized lunch status, parental education, home possessions (separate indicator variables for computer and books) as well as grade and subject fixed effects. Additional policy controls (state-by-year level, unless otherwise stated): Model 1 controls for mandatory state approval, where state has discretion, licensing, or accreditation of private schools; Model 2 controls for mandatory state licensure of private school teachers; Model 3 controls for mandatory registration or licensing of private schools (note: if approval is required, registration is also coded as being required); Model 4 controls for extent of private school curriculum control; Model 5 controls for tax credit/deduction law for scholarship contributions or educational expenses of parents; Model 6 controls for publicly funded voucher laws. Standard errors clustered at the state level. Regressions use population weights and standard errors clustered at the state level. Single, double, and triple asterisks indicate statistical significance at the 10%, 5%, and 1% levels, respectively. Data source: Sorens et al. (2008); Jordan and Grossmann (2020); See Figure 2

Table A.8 – Effect of CCSS exposure on student achievement in non-targeted subjects, robustness with additional controls for homeschooling policies

	Control for:							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Curriculum	Statute	Notice Extent	Notice Frequency	Notice Index	Recordkeeping	Testing	Teachers
CCSS Exposure	-0.082** (0.036)	-0.080** (0.036)	-0.079** (0.036)	-0.086** (0.036)	-0.089** (0.036)	-0.080** (0.036)	-0.079** (0.036)	-0.080** (0.035)
State and Year FEs	YES	YES	YES	YES	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES	YES	YES	YES
Adj. R-squared	0.390	0.390	0.390	0.390	0.391	0.390	0.390	0.390
Observations	1,099,880	1,099,880	1,099,880	1,099,880	1,099,880	1,099,880	1,099,880	1,099,880

Note: Each entry is from a separate two-way fixed effects regression model. Dependent variable: Standardized student achievement in subjects not targeted by the CCSS (Pool of science, civics, economics, geography, and history). Explanatory variable: Share of schooling years a student was exposed to CCSS (at the time of testing). Controls: Indicator variables for gender, races/ethnicities, English language learner status, disability status, subsidized lunch status, parental education, home possessions (separate indicator variables for computer and books) as well as grade and subject fixed effects. Additional policy controls (state-by-year level, unless otherwise stated): Model 1 controls for subjects/curriculum requirement for homeschoolers; Model 2 controls for whether homeschooling is explicitly permitted by statute; Model 3 controls for extent of homeschooling notice requirement; Model 4 controls for frequency of homeschooling notice requirement; Model 5 controls for homeschooling notification index (Extent of homeschooling notice requirement * Frequency of homeschooling notice requirement); Model 6 controls for extent of homeschool recordkeeping requirements; Model 7 controls for standardized testing or other official evaluation requirement of homeschooling; Model 8 controls for homeschooling teacher qualifications requirement. Regressions use population weights and standard errors clustered at the state level. Single, double, and triple asterisks indicate statistical significance at the 10%, 5%, and 1% levels, respectively. Data source: Sorens et al. (2008); Jordan and Grossmann (2020); See Figure 2

Table A.9 – Effect of CCSS exposure on student achievement in non-targeted subjects, robustness with additional controls for compulsory schooling policies

	Control for:			
	(1) Compulsory school age, lower bound	(2) Compulsory school age, upper bound	(3) Compulsory school years	(4) Kindergarten attendance
CCSS Exposure	-0.079** (0.037)	-0.079** (0.035)	-0.081** (0.036)	-0.082** (0.036)
State and Year FEs	YES	YES	YES	YES
Controls	YES	YES	YES	YES
Adj. R-squared	0.390	0.390	0.390	0.390
Observations	1,099,880	1,099,880	1,099,880	1,099,880

Note: Each entry is from a separate two-way fixed effects regression model. Dependent variable: Standardized student achievement in subjects not targeted by the CCSS (Pool of science, civics, economics, geography, and history). Explanatory variable: Share of schooling years a student was exposed to CCSS (at the time of testing). Controls: Indicator variables for gender, races/ethnicities, English language learner status, disability status, subsidized lunch status, parental education, home possessions (separate indicator variables for computer and books) as well as grade and subject fixed effects. Additional policy controls (state-by-year level, unless otherwise stated): Model 1 controls for compulsory school age, lower bound (minimum standard if set by local school district; age at which parental waivers not permitted); Model 2 controls for compulsory school age, upper bound (minimum standard if set by local school district; age at which parental waivers not permitted); Model 3 controls for compulsory school years (Compulsory school age, upper bound – Compulsory school age, lower bound); Model 4 controls for kindergarten attendance requirement. Regressions use population weights and standard errors clustered at the state level. Single, double, and triple asterisks indicate statistical significance at the 10%, 5%, and 1% levels, respectively. Data source: Sorens et al. (2008); Jordan and Grossmann (2020); See Figure 2

Table A.10 – Effect of CCSS exposure on student achievement in non-targeted subjects, further robustness checks

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Exclude already-treated states from controls		Controls: Governor Party	Controls: State Legislature	Controls: Governor Party and State Legislature	Controls: No parental education	Controls: No parental education & No imputation of missings	Sample: No Charter Schools	No weights
CCSS Exposure	-0.082** (0.036)	-0.082** (0.037)	-0.081** (0.037)	-0.084** (0.038)	-0.079** (0.036)	-0.090** (0.036)	-0.079** (0.036)	-0.093*** (0.025)
State and Year FEs	YES	YES	YES	YES	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES	YES	YES	YES
Adj. R-squared	0.392	0.391	0.390	0.391	0.380	0.381	0.390	0.399
Observations	996,390	1,060,780	1,099,880	1,060,780	1,103,630	1,067,950	1,077,420	1,103,630

Note: Each entry is from a separate two-way fixed effects regression model. Dependent variable: Standardized student achievement in subjects not targeted by the CCSS (Pool of science, civics, economics, geography, and history). Explanatory variable: Share of schooling years a student was exposed to CCSS (at the time of testing). Controls: Indicator variables for gender, races/ethnicities, English language learner status, disability status, subsidized lunch status, parental education, home possessions (separate indicator variables for computer and books) as well as grade and subject fixed effects. Regressions use population weights and standard errors clustered at the state level. Model 1 excludes states which adopted the CCSS in 2011 and 2012 from the sample which implies that no students from already-treated states act as controls; Model 2 controls for party of governor (equals one if Republican, and zero otherwise); Model 3 controls for party of state legislative chambers (equals one if both chambers Republican, and zero otherwise); Model 4 controls for party of governor and party of state legislative chambers; Model 5 excludes parental education from set of control variables; Model 6 excludes parental education from set of control variables and does not impute other missing control variables; Model 7 excludes charter schools from the sample of public schools; Model 8 does not use population weights. Single, double, and triple asterisks indicate statistical significance at the 10%, 5%, and 1% levels, respectively. Data source: See Figure 2

B ANALYSIS OF THE EFFECTS OF THE CCSS ON TARGETED SUBJECTS

We show evidence that the CCSS had, at best, modestly positive effects on student achievement in the targeted subjects math and ELA. This analysis largely confirms the conclusions Bleiberg (2021) has drawn on this question, although our findings (using more data, among other conceptual differences) suggest that Bleiberg (2021) rather overestimates than underestimates the positive effects of the CCSS on student achievement in targeted subjects.

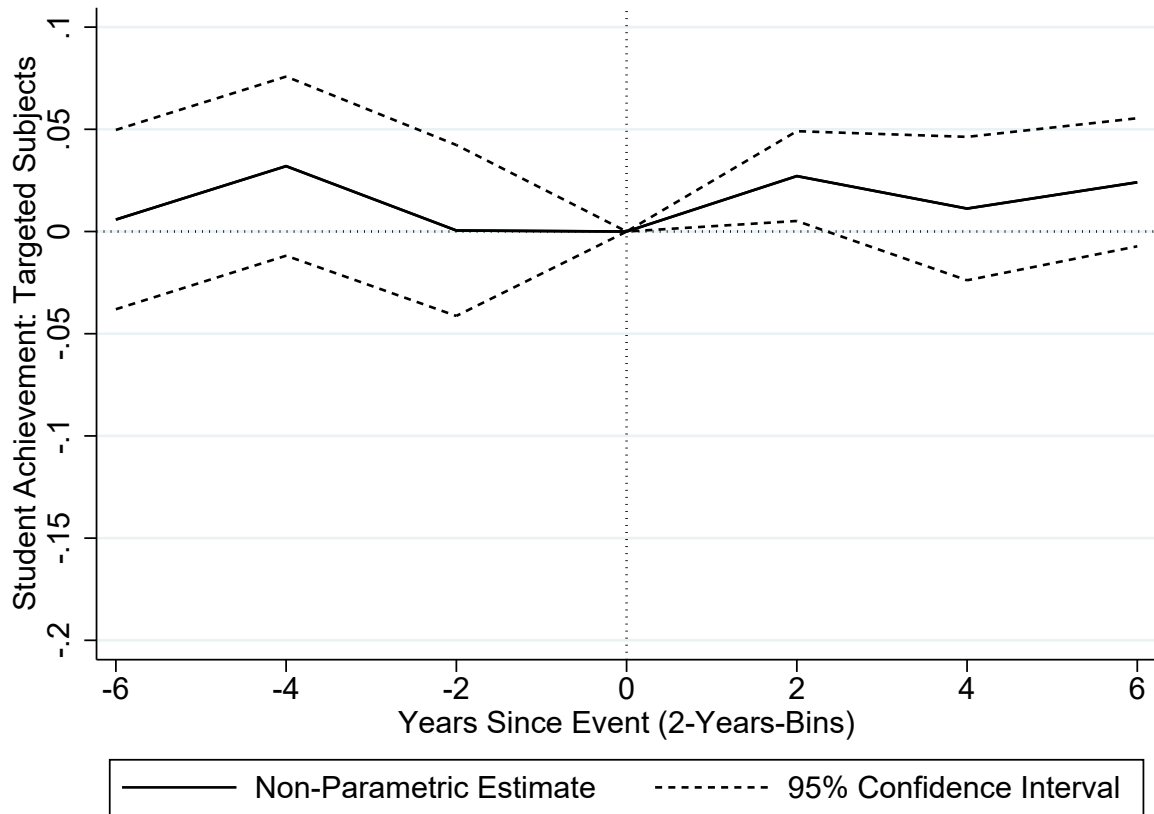
First, we visualize the modest positive effects in an event-study graph depicted in Figure B.1. The estimation equation follows Equation 2 presented in Section 3, with T_{istuv} now pooling standardized student achievement in all subjects that are targeted by the CCSS, across all available grades.¹⁹ The event-study graph shows a modest increase in student achievement after the adoption of the CCSS that is marginally significant directly after adoption and insignificantly different from zero thereafter. Using a second dataset on student achievement from the Stanford Education Data Archive SEDA 4.0 (Reardon et al., 2021) with a shorter pre- but longer post-period relative to the NAEP, we find basically null effects (see Figure B.2). SEDA does not contain data on non-targeted subjects which is why we cannot use it for the main analysis.

Second, we also show parametric two-way fixed effects DD results. The estimation equation follows Equation 1 presented in Section 3, with T_{istuv} pooling standardized student achievement in all subjects that are targeted by the CCSS, across all available grades (as above). As reported in Table B.11, the baseline specification and the econometric robustness specifications including state-specific time trend controls and the DDD model yield zero to small positive point estimates. Next, we analyse robustness regarding the treatment indicator as described in Section 6 and Appendix C. As shown

¹⁹Table B.13 provides the list of grades in which the NAEP tests were administered for each of the four targeted subjects math, reading, writing, and vocabulary (and for which state identifiers of students are available). We use the student-level data from all these subject-year-grade combinations in our analysis.

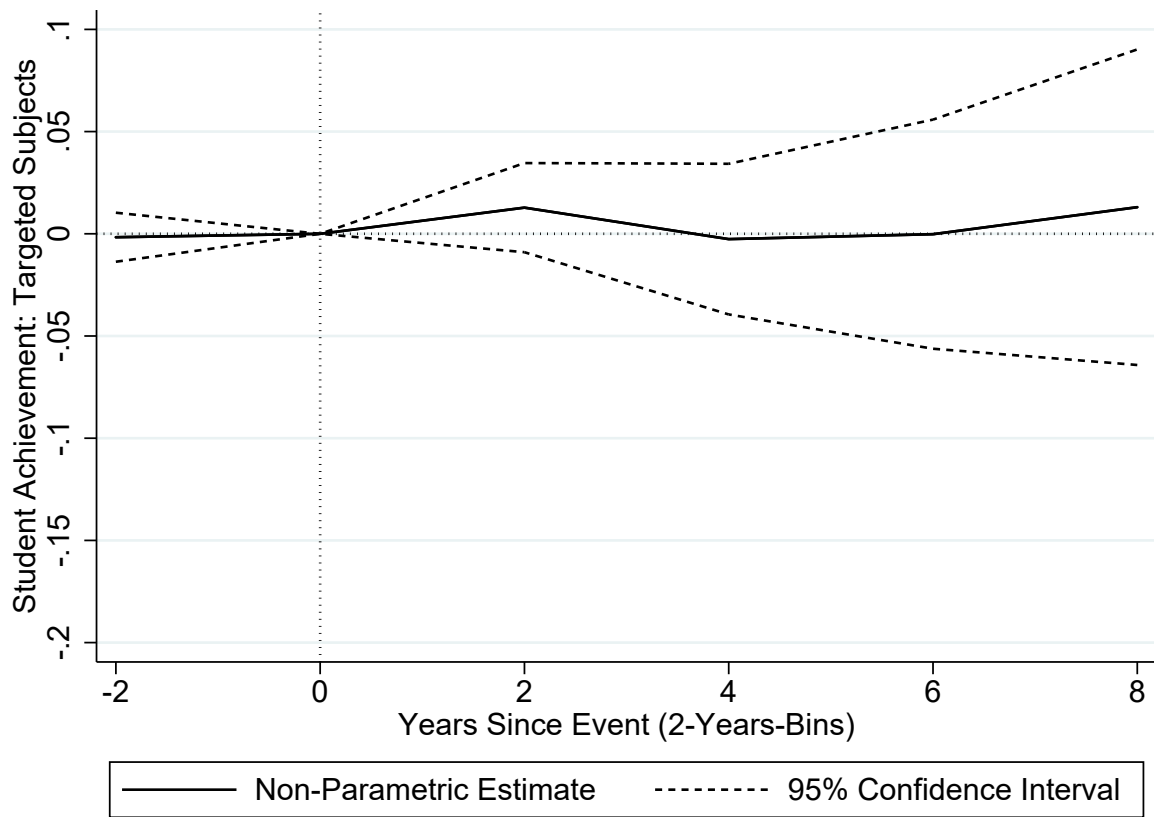
in Table B.12, we find zero to modestly positive effects across specifications.

Figure B.1 – Event-study graph: Targeted subjects (NAEP)



Note: Coefficients from non-parametric event-study regressions and their 95% confidence intervals. Dependent variable: Standardized student achievement in subjects targeted by the CCSS (Pool of math, reading, vocabulary and writing). Controls: Indicator variables for gender, races/ethnicities, English language learner status, disability status, subsidized lunch status, parental education, home possessions (separate indicator variables for computer and books) as well as state, test year, grade and subject fixed effects. Regressions use population weights and standard errors clustered at the state level. Numbers on horizontal axis refer to respective two-year bins; i.e. 2 = first two years of treatment (year 0 = excluded category). The p values of omnibus hypothesis tests of zero pre- and post-event effects are 0.003 and 0.075, respectively. Data source: See Figure 2

Figure B.2 – Event-study graph: Targeted subjects (SEDA)



Note: Coefficients from non-parametric event-study regressions and their 95% confidence intervals. Dependent variable: Standardized student achievement in subjects targeted by the CCSS (Pool of math and ELA). Controls: District shares of races/ethnicities, English language learner status, disability status, subsidized lunch status, economic disadvantage, rural location, as well as state, test year, grade and subject fixed effects. Regressions use precision weights (the inverse of the standard error of average student achievement in math and ELA squared) and standard errors clustered at the state level. Numbers on horizontal axis refer to respective two-year bins; i.e. 2 = first two years of treatment (year 0 = excluded category). The p values of omnibus hypothesis tests of zero pre- and post-event effects are 0.782 and 0.190, respectively. Data source: Reardon et al. (2021)

Table B.11 – Effect of CCSS exposure on student achievement in targeted subjects, econometric robustness

	(1)	(2)	(3)	(4)	(5)	(6)
	DD	(1) + linear state trends	(2) + quadratic state trends	(3) + cubic state trends	DDD	DD with private schools
CCSS Exposure	0.010 (0.023)	0.004 (0.029)	0.002 (0.051)	-0.001 (0.051)	0.007 (0.055)	0.011 (0.020)
State and Year FEs	YES	YES	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES	YES
Adj. R-squared	0.352	0.353	0.354	0.354	0.353	0.349
Observations	6,392,940	6,392,940	6,392,940	6,392,940	6,550,510	6,550,510

Note: Each entry is from a separate two-way fixed effects regression model, where Model (1) is the baseline model, Models (2), (3), (4) subsequently add linear, quadratic, and cubic state-specific time trends, and Model (5) presents a triple-difference model where school type (public vs. private school students), school type*state, school type*year, and state*year fixed effects replace all state-specific time trends. Model (6) estimates the basic two-way fixed effects model on a sample of public and private school students. Dependent variable: Standardized student achievement in subjects targeted by the CCSS (Pool of math, reading, vocabulary and writing). Explanatory variable: Share of schooling years a student was exposed to CCSS (at the time of testing). Controls: Indicator variables for gender, races/ethnicities, English language learner status, disability status, subsidized lunch status, parental education, home possessions (separate indicator variables for computer and books) as well as grade and subject fixed effects. Regressions use population weights and standard errors clustered at the state level. Single, double, and triple asterisks indicate statistical significance at the 10%, 5%, and 1% levels, respectively. Data source: See Figure 2

Table B.12 – Effect of CCSS exposure on student achievement in targeted subjects using different definitions of treatment implementation

	CCSS adoption	CCSS implementation requirement	CCSS implementation strategies	Effective CCSS implementation	Include temporary CCSS adopters and implementers	CCSS-aligned testing
	(1)	(2)	(3)	(4)	(5)	(6)
CCSS Exposure	0.010 (0.023)	-0.002 (0.027)	0.019 (0.019)	-0.001 (0.021)	0.046*** (0.017)	-0.007 (0.032)
State and Year FEs	YES	YES	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES	YES
Adj. R-squared	0.352	0.352	0.352	0.352	0.352	0.352
Observations	6,392,940	6,392,940	6,392,940	6,392,940	6,392,940	6,392,940

Note: Each entry is from a separate two-way fixed effects regression model. Dependent variable: Standardized student achievement in subjects targeted by the CCSS (Pool of math, reading, vocabulary and writing). Explanatory variables: Share of schooling years a student was exposed to CCSS (at the time of testing), where in Models 1 (CCSS adoption, baseline model) each schooling year counts as exposed in a given state in which the state adopted the CCSS permanently before that year or in the same year according to Achieve Inc. (2013) and CCSSI (2023), where Model 2 (CCSS implementation requirement) each schooling year counts as exposed in a given state in which the state expects teachers to fully incorporate CCSS into classroom instruction in grades K-12 in English language arts and mathematics according to Achieve Inc. (2013) and CCSSI (2023), where in Models 3 (CCSS implementation strategies) each schooling year counts as exposed in a given state if state education agency officials report that their state pursued at least two out of three CCSS implementation strategies (professional development, new instructional materials, joined testing consortium) as reported in Webber et al. (2014), where in Models 4 (Effective CCSS implementation) each schooling year counts as exposed in a given state in which the state implemented an effective change in state standard content through the adoption of CCSS which we define as not having had a state standard in place before the adoption of CCSS whose academic rigor is "too close to call" in comparison with CCSS (Carmichael et al., 2010) for the set of states adopting CCSS according to Achieve Inc. (2013) and CCSSI (2023), where in Models 5 (Include temporary CCSS adopters) each schooling year counts as exposed in a given state in which a state adopted and/or implemented CCSS at least temporarily according to Bleiberg (2021); and where in Models 6 (CCSS-aligned testing) each schooling year counts as exposed in a given state in which the state adopted CCSS-aligned standardized testing including field and transitional tests according to our own research (see Table C.14 for state-specific details of CCSS-aligned testing). Table C.15 provides state-specific coding information on all treatment definitions. Controls: Indicator variables for gender, races/ethnicities, English language learner status, disability status, subsidized lunch status, parental education, home possessions (separate indicator variables for computer and books) as well as grade and subject fixed effects. Regressions use population weights and standard errors clustered at the state level. Single, double, and triple asterisks indicate statistical significance at the 10%, 5%, and 1% levels, respectively. Data source: See Figure 2

Table B.13 – List of grades of NAEP tests for targeted subjects

Year	Targeted Subjects			
	Math	Reading	Vocabulary	Writing
2002		4, 8, 12		8, 12
2003	4, 8	4, 8		
2004				
2005	4, 8, 12	4, 8, 12		4, 8, 12
2006				
2007	4, 8	4, 8		8, 12
2008				
2009	4, 8, 12	4, 8, 12	4, 8, 12	
2010				
2011	4, 8	4, 8	4, 8	8, 12
2012				
2013	4, 8	4, 8		
2014				
2015	4, 8, 12	4, 8, 12		
2016				
2017	4, 8,	4, 8		

Note: NAEP student achievement data in targeted subjects at the subject-by-year-by-grade level. Data source: See [Figure 2](#)

C BACKGROUND INFORMATION ON TREATMENT

DEFINITION ROBUSTNESS

To test whether our results hold if we define treatment based on CCSS implementation, we re-run our main regression using five different treatment variables each capturing different information about CCSS implementation in Section 6. This appendix provides background information about the construction and data sources of these five alternative treatment definitions.

First, we collect information on CCSS implementation requirements, from Achieve Inc. (2013) and CCSSI (2023). Here, the year of full implementation of CCSS is defined as the school year the respective state expects teachers in grades K-12 in math and ELA to incorporate the standards into classroom instruction. The time between adoption and full implementation varies between 1 to 4 years across adopting states, with an average of about 3 years.

Second, we note that state expectations about teachers implementing the CCSS into classroom instruction do not necessarily have to be aligned with actual state efforts to implement the CCSS. However, the latter might be more relevant for ultimate exposure of students to the CCSS and potential effects on student achievement than formal state expectations. To incorporate this idea into our analysis, we make use of a survey of state education agency officials provided by Webber et al. (2014). They conducted a survey of state education agency officials which collects information on actual state efforts towards CCSS implementation. Specifically, the survey respondents answer questions about whether the state has provided, guided or funded professional development on the CCSS, whether it has provided curriculum or instructional materials for the CCSS, and whether it has worked with a federally funded consortium to develop assessments aligned with the CCSS. In this treatment coding, we count a schooling year in a given state as being exposed to the CCSS if this state has adopted the CCSS according to Achieve Inc. (2013) and CCSSI (2023) and pursued at least two out of three CCSS implementation

strategies as reported by the state education agency officials.

Third, we calculate a treatment indicator capturing effective CCSS implementation. Here, we build on the idea that effective change of the state standard can only be induced by the CCSS if the state standard in place prior to the adoption of the CCSS in the state in question is sufficiently different from the CCSS. To this end, we make use of a comparison of academic rigor of the CCSS with the respective state standards in place prior to the CCSS provided by Carmichael et al. (2010). We code students from states as being in the control group at all years if their pre-CCSS state standards are “too close to call” in both math and ELA in a comparison with the CCSS (in addition to coding students from states that did not adopt the CCSS according to Achieve Inc. (2013) and CCSSI (2023) as being in the control group at all years).

Fourth, we account for the fact that some states may have adopted and/or implemented some elements of the CCSS temporarily, even when they are listed as non-adopters and non-implementers in Achieve Inc. (2013) and CCSSI (2023). According to Bleiberg (2021), four of the eight non-permanent adopters of the CCSS in the coding based on Achieve Inc. (2013) and CCSSI (2023) have implemented at least some elements of the CCSS temporarily. The map presented in Figure 1 depicts them as temporary adopters. In this treatment coding, we count a schooling year as being exposed to the CCSS if the state in question adopted the CCSS temporarily or permanently.

Fifth, we argue that the relevant criterion for actual CCSS implementation might be the alignment of the content of state-mandated standardized testing with the CCSS. To assess this hypothesis, we did own background research to find out which state mandated what type of standardized test for each grade group and year. State-specific details on which tests (including field and transitional tests) are mandated when and for which grade are reported in Table C.14. Subsequently, we assessed which of these tests are aligned with the CCSS. This analysis allowed us to infer the year in which CCSS-aligned standardized testing was mandated in a given state. In the corresponding treatment coding, we count a schooling year in a given state as being exposed to the CCSS if this state has mandated

CCSS-aligned standardized testing in any group grade in that year.

Table C.15 presents the treatment status for each state for the baseline definition of CCSS adoption and the five definitions of CCSS implementation. In particular, it shows whether schooling years in a given state never count as being exposed to the CCSS ("always control"), or, if they do, from which year onwards.

Table C.14 – State-mandated tests, by state and grade group from 2010 onwards (based on own research)

State	3-8 grades	High school
Alabama	2010: Alabama Reading and Math Test (ARMT+) 2014: ACT Aspire	2010: Alabama High School Graduation Exam (AHSGE) 2014: ACT End of course
Alaska	2010: Standards-Based Assessments (SBAs) 2015: Alaska Measures of Progress (AMP) 2017: Performance Evaluation for Alaska's Schools (PEAKS)	2010: SBAs 2015: AMP; ACT, SAT, or WorkKeys 2017: PEAKS; ACT, SAT, or WorkKeys
Arizona	2010: Arizona Instrument to Measure Standards (AIMS) 2014: Field test PARCC 2015: AzMerit	2010: AIMS 2014: Field test PARCC 2015: AzMerit
Arkansas	2010: Arkansas Comprehensive Testing, Assessment, and Accountability Program (ACTAAP) 2013: Arkansas Benchmark 2014: Field test PARCC 2015: PARCC 2016: ACT Aspire	2010: Arkansas Comprehensive Testing, Assessment, and Accountability Program (ACTAAP) 2013: Arkansas Benchmark 2014: Field test PARCC 2015: PARCC 2016: ACT Aspire
California	2010: Standardized Testing and Reporting (STAR) 2014: Field test Smarter Balanced 2015: Smarter Balanced	2010: Standardized Testing and Reporting (STAR) 2014: Field test Smarter Balanced 2015: Smarter Balanced

Table C.14 - Aligned Testing Details (continued)

State	3-8 grades	High school
Colorado	2010: Colorado Student Assessment Program (CSAP) 2012: Transitional Colorado Assessment Program (TCAP) 2014: Field test PARCC 2015: PARCC	2010: Colorado Student Assessment Program (CSAP) 2012: Transitional Colorado Assessment Program (TCAP) 2014: Field test PARCC 2015: PARCC 2016: PSAT, ACT 2017: SAT
Connecticut	2010: Connecticut Mastery Test (CMT) 2014: Field test Smarter Balanced 2015: Smarter Balanced	2010: Connecticut Academic Performance Test (CAPT) 2014: Field test Smarter Balanced 2015: Smarter Balanced 2016: SAT
Delaware	2010: Delaware Comprehensive Assessment System (DCAS) 2014: Field test Smarter Balanced 2015: Smarter Balanced	2010: Delaware Comprehensive Assessment System (DCAS) 2014: Field test Smarter Balanced 2015: Smarter Balanced 2016: SAT
District of Columbia	2010: District of Columbia Comprehensive Assessment System (DC CAS) 2012: DC CAS revised (transitional test) 2013: DC CAS revised 2014: Field test PARCC 2015: PARCC	2010: District of Columbia Comprehensive Assessment System (DC CAS) 2012: DC CAS revised (transitional test) 2013: DC CAS revised 2014: Field test PARCC 2015: PARCC
Florida	2010: Florida Comprehensive Assessment Test (FCAT) 2011: FCAT 2.0 2014: Florida Standards Assessment (FSA)	2010: FCAT 2011: Florida End-of-Course (EOC) Assessments 2014: FSA or Next Generation Sunshine State Standards (NGSSS) 2016: FSA

Table C.14 - Aligned Testing Details (continued)

State	3-8 grades	High school
Georgia	2010: Criterion-Referenced Competency Tests (CRCT) 2015: Georgia Milestones Assessment System (GMAS)	2010: End of Course Test (EOCT) 2015: GMAS
Hawaii	2010: Hawaii State Assessment (HSA) 2014: Part-HAS Part-Smarter Balanced (transition test) 2014: Field test Smarter Balanced 2015: Smarter Balanced	2010: Hawaii State Assessment (HSA) 2014: Part-HAS Part-Smarter Balanced (transition test) 2014: Field test Smarter Balanced 2015: Smarter Balanced
Idaho	2010: Idaho Standards Achievement Test (ISAT) 2013-14: Field test Smarter Balanced 2015: Smarter Balanced	2010: Idaho Standards Achievement Test (ISAT) 2013-14: Field test Smarter Balanced 2015: Smarter Balanced in 10th grade; Choice of ACT, SAT or ACT Compass for 11th grade.
Illinois	2010: Illinois Standards Achievement Tests (ISAT) 2014: Field test PARCC 2015: PARCC	2014: Field test PARCC 2015: PARCC 2016: SAT
Indiana	2010: Indiana Statewide Testing for Educational Progress Plus (ISTEP+)	2010: ISTEP+, end-of-course tests
Iowa	2010: Iowa Test of Basic Skills (ITBS) 2011: Iowa Assessments 2014: Field test Smarter Balanced	2010: Iowa Test of Educational Development (ITED) 2011: Iowa Assessments 2014: Field test Smarter Balanced
Kansas	2014: No test 2015: Field test Kansas State Assessment (KSA) 2016: KSA	2014: No test 2015: Field test Kansas State Assessment (KSA) 2016: KSA
Kentucky	2010: Kentucky Performance Rating for Educational Progress (K-PREP)	2010: K-PREP, ACT QualityCore, ACT

Table C.14 - Aligned Testing Details (continued)

State	3-8 grades	High school
Louisiana	<p>2006: Louisiana Educational Assessment Program (LEAP) and iLEAP</p> <p>2013: LEAP and iLEAP revised (transitional test)</p> <p>2014: Field test PARCC</p> <p>2015: PARCC</p> <p>2016: Mix of PARCC and LEAP</p>	<p>2010: End-of-course tests, ACT, ACT Plan</p> <p>2013-14: End-of-course revised (transitional test)</p> <p>2015: End-of-course revised</p>
Maine	<p>2010: New England Common Assessment Program (NECAP)</p> <p>2014: Field test Smarter Balanced</p> <p>2015: Smarter Balanced</p> <p>2016: Maine Educational Assessments (MEA)</p>	<p>2010: SAT</p> <p>2014: Field test Smarter Balanced</p> <p>2015: Smarter Balanced</p> <p>2015: SAT</p>
Maryland	<p>2010: Maryland State Assessment (MSA)</p> <p>2014: Field test PARCC</p> <p>2015: PARCC</p>	<p>2010: Maryland High School Assessment (HSA)</p> <p>2014: Field test PARCC</p> <p>2015: PARCC</p>
Massachusetts	<p>2010: Massachusetts Comprehensive Assessment System (MCAS)</p> <p>2014: Field test PARCC</p> <p>2014: Districts choose between PARCC or MCAS</p> <p>2016: Mix of PARCC and Next Generation MCAS</p> <p>2017: Next Generation MCAS</p>	<p>2010: Massachusetts Comprehensive Assessment System (MCAS)</p> <p>2014: Field test PARCC</p> <p>2014: Districts choose between PARCC or MCAS</p> <p>2015: MCAS</p>
Michigan	<p>2010: Michigan Educational Assessment Program (MEAP)</p> <p>2014: Field test Smarter Balanced</p> <p>2015: Michigan Student Test of Educational Progress (M-STEP)</p>	<p>2010: Michigan Merit Exam (MME: includes SAT, WorkKeys)</p> <p>2015: MME, PSAT</p>

Table C.14 - Aligned Testing Details (continued)

State	3-8 grades	High school
Minnesota	2010: Minnesota Comprehensive Assessments (MCA)	2010: MCA
Mississippi	2010: Mississippi Curriculum Test (MCT) 2014: Field test PARCC 2015: PARCC 2016: Mississippi Academic Assessment Program (MAAP)	2010: Subject Area Testing Program (SATP) 2014: Field test PARCC 2015: PARCC 2016: ACT
Missouri	2010: Missouri Assessment Program (MAP) 2014: MAP revised (transitional test) 2014: Field test Smarter Balanced 2015: Smarter Balanced 2016: MAP	2010: Missouri End-of-Course Assessments 2014: End of Course (EOC) revised (transitional test) 2015: EOC revised, ACT
Montana	2010: Montana’s Criterion Reference Test (Montana’s CRT) 2014: Field test Smarter Balanced 2015: Smarter Balanced	2010: Montana’s CRT 2014: Field test Smarter Balanced 2015: Smarter Balanced 2016: ACT
Nebraska	2010: Nebraska State Accountability Tests (NeSA)	2010: NeSA 2017: ACT
Nevada	2010: Nevada’s Criterion Reference Test (Nevada’s CRT) 2014: Field test Smarter Balanced 2015: Smarter Balanced	2010: High School Proficiency Examination (HSPE) 2015: ACT
New Hampshire	2010: New England Common Assessment Program (NECAP) 2013: NECAP revised (transitional test) 2014: Field test Smarter Balanced 2015: Smarter Balanced	2010: NECAP 2013: NECAP revised (transitional test) 2014: Field test Smarter Balanced 2015: Smarter Balanced 2016: PACE, SAT

Table C.14 - Aligned Testing Details (continued)

State	3-8 grades	High school
New Jersey	2010: New Jersey Assessment of Skills and Knowledge (NJASK) 2014: NJASK revised (transitional test) 2014: Field test PARCC 2015: PARCC	2010: High School Proficiency Assessment 2014: Field test PARCC 2015: PARCC
New Mexico	2010: New Mexico Standards-based Assessment (NMSBA) 2014: Field test PARCC 2015: PARCC	2010: NMSBA 2014: Field test PARCC 2015: PARCC
New York	2012: Field test 2013: New York State English Language Arts and Mathematics Tests 2014: Field test PARCC 2016: New York State Assessments	2013: Regents Exams 2014: Regents revised 2014: Field test PARCC
North Carolina	2012: Field test 2013: End-of-grade tests 2014: Field test Smarter Balanced	2012: Field test 2013: End-of-course tests, ACT PLAN, ACT, WorkKeys 2014: Field test Smarter Balanced
North Dakota	2014: Field test Smarter Balanced 2015: Smarter Balanced	2014: Field test Smarter Balanced 2015: Smarter Balanced
Ohio	2014: Ohio Achievement Assessments 2014: Field Test PARCC 2015: PARCC 2016: Ohio State Tests (OST)	2014: Ohio Graduation Tests 2014: Field Test PARCC 2015: PARCC 2016: OST, Ohio Graduation Test 2017: End-of-course tests, SAT/ACT
Oklahoma	2010: Oklahoma Core Curriculum Test (OCCT) 2017: Oklahoma School Testing Program (OSTP)	2010: End-of-course tests 2017: OSTP

Table C.14 - Aligned Testing Details (continued)

State	3-8 grades	High school
Oregon	2010: Oregon Assessment of Knowledge and Skills 2014: Field test Smarter Balanced 2015: Smarter Balanced	2010: Oregon Assessment of Knowledge and Skills 2014: Field test Smarter Balanced 2015: Smarter Balanced
Pennsylvania	2010: Pennsylvania System of School Assessment (PSSA) 2013: Field test PSSA revised 2015: PSSA revised	2010: Pennsylvania System of School Assessment (PSSA) 2013: Keystone Exams
Rhode Island	2010: New England Common Assessment Program (NECAP) 2014: Field test PARCC 2015: PARCC	2010: New England Common Assessment Program (NECAP) 2014: Field test PARCC 2015: PARCC
South Carolina	2010: South Carolina Palmetto Assessment of State Standards (SCPASS) 2015: ACT Aspire 2016: SC Ready	2014: ACT Plus Writing, ACT WorkKeys 2015: End-of-course tests, ACT
South Dakota	2014: Field test Smarter Balanced 2015: Smarter Balanced	2014: Field test Smarter Balanced 2015: Smarter Balanced
Tennessee	2014: Tennessee Comprehensive Assessment Program (TCAP) 2014: Field test PARCC 2015: TNReady	2014: Tennessee Comprehensive Assessment Program (TCAP) 2014: Field test PARCC 2015: TNReady
Texas	2015: State of Texas Assessments of Academic Readiness (STAAR)	2015: STARR
Utah	2014: Field test Student Assessment of Growth and Excellence (SAGE) 2015: SAGE	2014: Field test Student Assessment of Growth and Excellence (SAGE) 2015: SAGE, ACT
Vermont	2010: New England Common Assessment Program (NECAP) 2014: Field test Smarter Balanced 2015: Smarter Balanced	2010: New England Common Assessment Program (NECAP) 2014: Field test Smarter Balanced 2015: Smarter Balanced

Table C.14 - Aligned Testing Details (continued)

State	3-8 grades	High school
Virginia	2010: Standards of Learning (SOL)	2010: SOL
Washington	2014: Field test Smarter Balanced 2015: Smarter Balanced	2010: High School Proficiency Exam 2014: Field test Smarter Balanced 2015: Smarter Balanced
West Virginia	2014: Field test Smarter Balanced 2015: Smarter Balanced	2014: Field test Smarter Balanced 2015: Smarter Balanced
Wisconsin	2010: Wisconsin Knowledge and Concepts Exam (WKCE) 2014: Field test Smarter Balanced 2015: Smarter Balanced 2016: Wisconsin Forward	2015: ACT, ACT Aspire
Wyoming	2010: Proficiency Assessments for Wyoming Students (PAWS), Student Assessment of Writing Skills (SAWS) 2013: Field test PAWS revised 2014: PAWS revised, Field test Smarter Balanced 2017: Wyoming Test of Proficiency and Progress (WY-TOPP)	2016: ACT Aspire (9-10), ACT (11) 2017: Wyoming Test of Proficiency and Progress (WY-TOPP), ACT

Table C.15 – Treatment codings by state

State	Main treatment: CCSS Adoption	CCSS implementation requirement	CCSS implementation strategies	Effective CCSS implementation	Include temporary CCSS adopters and implementers	CC-aligned testing
Alabama	2010	2013	2010	Always Control	2010	2014
Alaska	Always Control	Always Control	Always Control	Always Control	Always Control	Always Control
Arizona	2010	2013	2010	2010	2010	2014
Arkansas	2010	2013	2010	2010	2010	2014
California	2010	2014	2010	2010	2010	2014
Colorado	2010	2013	2010	2010	2010	2012
Connecticut	2010	2013	2010	2010	2010	2014
Delaware	2010	2012	2010	2010	2010	2014
D.C.	2010	2012	Always Control	2010	2010	2012
Florida	Always Control	Always Control	2010	Always Control	2010	Always Control
Georgia	2010	2014	2010	Always Control	2010	2015
Hawaii	2010	2013	2010	2010	2010	2014
Idaho	2011	2013	2010	2011	2011	2013
Illinois	2010	2013	2010	2010	2010	2014
Indiana	Always Control	Always Control	2010	Always Control	2010	Always Control
Iowa	2010	2012	2010	2010	2010	2011
Kansas	2010	2013	2010	2010	2010	2015
Kentucky	2010	2011	2010	2010	2010	2010
Louisiana	2010	2013	Always Control	2010	2010	2013
Maine	2011	2012	2010	2011	2011	2010
Maryland	2010	2013	2010	2010	2010	2014
Massachusetts	2010	2013	2010	Always Control	2010	2010
Michigan	2010	2012	2010	2010	2010	2010
Minnesota	2010	2012	Always Control	2010	2010	Always Control
Mississippi	2010	2013	2010	2010	2010	2014
Missouri	2010	2014	2010	2010	2010	2014
Montana	2011	2013	Always Control	2011	2011	2014

Treatment codings by state (continued)

State	Main treatment: CCSS Adoption	CCSS implementation requirement	CCSS implementation strategies	Effective CCSS implementation	Include temporary CCSS adopters and implementers	CC-aligned testing
Nebraska	Always Control	Always Control	Always Control	Always Control	Always Control	Always Control
Nevada	2010	2013	2010	2010	2010	2014
New Hampshire	2010	2014	2010	2010	2010	2013
New Jersey	2010	2013	2010	2010	2010	2014
New Mexico	2010	2013	2010	2010	2010	2014
New York	2010	2013	2010	2010	2010	2012
North Carolina	2010	2012	2010	2010	2010	2012
North Dakota	2011	2013	2010	2011	2010	2015
Ohio	2010	2013	2010	2010	2010	2014
Oklahoma	Always Control	Always Control	2010	Always Control	2010	Always Control
Oregon	2010	2014	2010	2010	2010	2014
Pennsylvania	2010	2013	2010	2010	2010	2013
Rhode Island	2010	2013	2010	2010	2010	2014
South Carolina	Always Control	Always Control	2010	Always Control	2010	Always Control
South Dakota	2010	2014	Always Control	2010	2010	2014
Tennessee	2010	2013	2010	2010	2010	2014
Texas	Always Control	Always Control	Always Control	Always Control	Always Control	Always Control
Utah	2010	2013	2010	2010	2010	2014
Vermont	2010	2013	2010	2010	2010	2014
Virginia	Always Control	Always Control	Always Control	Always Control	Always Control	Always Control
Washington	2011	2014	Always Control	2011	2012	2014
West Virginia	2010	2014	2010	2010	2010	2014
Wisconsin	2010	2014	2010	2010	2010	2014
Wyoming	2012	2014	Always Control	2012	2012	2013

Note: Table shows whether schooling years in a given state are never coded as exposed to CCSS ("always control"), or, if they are, from which year onwards, for different treatment definitions. Data sources: Achieve Inc. (2013) and CCSSI (2023) (CCSS Adoption and CCSS implementation requirements); Webber et al. (2014) (CCSS implementation strategies); Carmichael et al. (2010) (Effective CCSS implementation); Bleiberg (2021) (Include temporary CCSS adopters and implementers); Own research, see also C.14 (CCSS-aligned testing)